

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Sleep Dynamics and Seizure Control in a Mesoscale Cortical Model

Permalink

<https://escholarship.org/uc/item/6r91v65z>

Author

Lopour, Beth Ann

Publication Date

2009

Peer reviewed|Thesis/dissertation

Sleep Dynamics and Seizure Control in a Mesoscale Cortical Model

by

Beth Ann Lopour

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 Szeri, Chair

Professor Alexandre Bayen

Professor J. Karl Hedrick

Professor Bruno Olshausen

Fall 2009

Sleep Dynamics and Seizure Control in a Mesoscale Cortical Model

©2009

by Beth Ann Lopour

Abstract

Sleep Dynamics and Seizure Control in a Mesoscale Cortical Model

by

Beth Ann Lopour

Doctor of Philosophy in Engineering - Mechanical Engineering

University of California, Berkeley

Professor Andrew Szeri, Chair

Using a mean-field, mesoscale model of the human cortex, we examine both feedback control for the suppression of seizures and the continuous mapping of sleep states to aid in seizure prediction. First, we verify the strong convergence of numerical solutions to the model, which consists of coupled stochastic partial differential equations. In doing this, we pay special attention to the sharp spatial changes that occur at electrode edges. This allows us to choose appropriate step sizes for our simulations; because the spatial step size must be small relative to the size of an electrode in order to resolve its electrical behavior, we are able to include a more detailed electrode profile in the simulation. We then develop a new model for the measurement of a cortical surface electrode based on extracellular currents flowing in the cortex. This model is used to simulate feedback with a new control algorithm that utilizes a charge-balanced signal. Not only does this succeed in suppressing the seizure oscillations, but it guarantees that the applied signal will be charge-balanced and therefore unlikely to cause cortical damage.

Next, we turn to a representation of the human sleep cycle contained within the mesoscale cortical model. We show that it can be used to obtain a detailed description of electroencephalogram (EEG) sleep stages, and we discuss how this analysis may aid in the prediction and prevention of seizures during sleep. The association between EEG data and the cortical model is found via locally linear embedding (LLE), a method of nonlinear dimensionality reduction. We first show that LLE can distinguish between traditional sleep stages when applied to EEG data. It reliably separates rapid eye movement (REM) and non-REM sleep and maps the EEG data to a low-dimensional output space where the sleep state changes smoothly over time. We also incorporate the concept of strongly connected components and use this as a method of automatic outlier rejection for EEG data. Then, by using LLE on a hybrid data set containing both sleep EEG and signals generated from the mesoscale cortical model, we quantify the relationship between the data and the mathematical model. This enables us to take any sample of sleep EEG data and associate it with a position among the continuous range of sleep states provided by the model; we can thus infer a trajectory of states as the subject sleeps. Because current sleep scoring consists of only five discrete stages, this technique may allow for tracking of sleep dynamics and a greater ability to predict the

onset of seizures. We show that this method gives consistent results for various subjects over a full night of sleep and can be done in real time.

To my family and friends, for their love, loyalty, and laughter.

Contents

List of Figures	iv
List of Tables	xii
1 Introduction	1
1.1 Seizures and epilepsy	1
1.1.1 Clinical presentation and diagnosis	1
1.1.2 Seizure onset during sleep and waking	2
1.1.3 Treatment	3
1.2 Computational Neuroscience, Epilepsy, and Sleep	4
1.2.1 Feedback control of epileptic seizures	5
1.2.2 A continuous mapping of human sleep states	5
1.3 Summary of Original Contributions	6
2 Mesoscale Cortical Model	7
2.1 Description of the model	7
2.2 Representations of cortical states in the model	10
2.2.1 Seizure waves	10
2.2.2 Sleep states	11
3 Mathematical methods	16
3.1 Strong convergence of numerical solutions	16
3.1.1 Method	16
3.1.2 Results	17
3.2 Data analysis techniques	19
3.2.1 Statistical analysis of EEG data	19
3.2.2 Locally Linear Embedding (LLE)	21
3.2.3 Identification of strongly connected components	23
4 Feedback control for the suppression of epileptic seizures	25
4.1 Refinements to the model of feedback control	25
4.1.1 Electrode profile	25
4.1.2 Model of electrode measurements	26
4.1.3 Charge-balanced control	30
4.2 Simulations of feedback control	32

4.2.1	Proportional control	32
4.2.2	Differential control	34
4.2.3	Total integral feedback control	36
4.3	Spatial properties of the feedback control model	37
4.3.1	The cortical length scale	37
4.3.2	Implications for electrode design	38
5	A continuous mapping of EEG sleep states	43
5.1	Sleep EEG	43
5.1.1	Traditional EEG sleep stages	43
5.1.2	PhysioBank data sets	44
5.2	LLE applied to sleep EEG data	45
5.2.1	LLE input based on EEG features	45
5.2.2	Automated ranking of feature sets	46
5.2.3	Separation of sleep stages via LLE	47
5.3	Integration of EEG data and the model sleep cycle	49
5.3.1	Model data set	50
5.3.2	Application of LLE to a hybrid data set	52
5.3.3	Connection to the theoretical sleep manifold	53
5.3.4	Inclusion of additional data sets	56
6	Discussion	59
6.1	Summary of contributions	59
6.2	Future Work	60
	Bibliography	62
A	Stochastic signals at relative step sizes	67
A.1	Generation of two-dimensional stochastic signals	67
A.2	Scaling preserves signal variance	69
B	Stochastic signals at absolute step sizes	71

List of Figures

- 2.1 Flow chart representation of model equations (2.1), (2.3), (2.5), and (2.7), which govern the excitatory population. The boxes describe the physiological significance of the model variables and parameters listed beneath them. Note that feedback occurs through the sigmoid function \tilde{S}_e , which is a function of \tilde{h}_e . This population is also coupled to the inhibitory population through local inputs described by $\tilde{S}_i[\tilde{h}_i]$. For a cell-based depiction of the model, see [54]. 9
- 2.2 (a) Typical behavior of the cortical model, with parameters $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$. Here we simulate a 200mm strip of cortex over a period of 0.5 seconds, and the value of h_e is represented by grayscale. For these parameters, the model exhibits low amplitude stochastic behavior. (b) Plot of h_e versus time extracted from Figure 2.2(a) at 100.8mm. 11
- 2.3 (a) Seizure waves traveling on the simulated cortex, with parameters $P_{ee} = 548.0$ and $\Gamma_e = 0.8 \times 10^{-3}$. The characteristics of the wave are determined by the distribution of P_{ee} ; here it is a Gaussian curve, so the wave starts in the center where P_{ee} is at its maximum (548.0) and propagates outward until the level of excitation is too low to support it. In this case, no feedback control is applied, so the waves will reoccur indefinitely. (b) Plot of h_e in time extracted from Figure 2.3(a) at 100.8mm. 12
- 2.4 The manifold of steady states in h_e from the mesoscale cortical model, hereafter referred to as the “sleep manifold.” The parameters L and Δh_e^{rest} represent the actions of adenosine and acetylcholine, neuromodulators that vary over the course of the human sleep cycle. The manifold has two stable solutions on its left side; a jump from the bottom solution to the top solution represents the fast transition between NREM and REM sleep. However, the slow transition from REM to NREM occurs smoothly down the right side of the manifold, where there is only one set of solutions. 13

- 3.1 Convergence of numerical solutions as the time step is decreased using the method described in section 3.1.1. Here the spatial terms in the model have been removed to reduce it to an ODE. We used $\Delta t_0 = 2.5 \times 10^{-4}$ seconds, and we plotted the solutions for $\Delta t = 8\Delta t_0$, $4\Delta t_0$, $2\Delta t_0$, and Δt_0 . The data is represented by open circles, x markers, solid dots, and a solid line, respectively. The two smallest time steps (the solid line and solid dots) give overlapping results, indicating that the solution has converged. Note that, although the points at $8\Delta t_0$ do not accurately capture the peaks and troughs in the data, their absolute position is accurate. This study was done with typical excitation $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$ 18
- 3.2 Convergence of numerical solutions over an electrode nonlinearity as the spatial step size is decreased using the method from section 3.1.1. The full simulation spanned 22.4mm, but here we show only 3mm of uncontrolled cortex and 2mm of cortex underneath an electrode while proportional feedback is applied. This figure shows solutions for $\Delta x = 4\Delta x_0$ (dashed), $2\Delta x_0$ (solid), and Δx_0 (solid), where $\Delta x_0 = 0.112\text{mm}$. Note that the two smallest time steps give very similar, overlapping solutions for \tilde{h}_e , while $4\Delta x_0$ gives an inaccurate result. Therefore, we choose the step size $\Delta x = 2\Delta x_0 = 0.224\text{mm}$ as a balance between accuracy and simulation cost. Other relevant parameters were $\Delta t = 2 \times 10^{-6}\text{s}$, $N = 80000$, $P_{ee} = 548.0$, and $\Gamma_e = 0.8 \times 10^{-3}$ 19
- 3.3 Convergence of numerical solutions as the spatial step size is decreased using the method described in section 3.1.1. Here we used $\Delta x_0 = 0.112\text{mm}$, and we plotted the solutions for $\Delta x = 4\Delta x_0$ (dashed), $2\Delta x_0$ (solid), and Δx_0 (solid). The two smallest time steps give very similar results, indicating that the solution has converged; this justifies our choice of step size, $\Delta x = 0.224\text{mm}$. This figure was created with typical excitation $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$, and $N = 80000$ time steps were used at $\Delta t = 2 \times 10^{-6}\text{s}$ 20
- 3.4 A simple example of LLE, where three dimensions are reduced to two. (a) The underlying manifold, which lives in 3D space but has only two dimensions. In a typical LLE problem, the shape of this manifold is unknown and has too many dimensions to visualize easily. (b) A sampling of points from the manifold, which serves as the input to the LLE algorithm. (c) The result of applying LLE to the data set in (b). Note that in the $Y_1 - Y_2$ output space, the manifold has been flattened to reveal its two principal dimensions. This figure was generated using the “scurve.m” code from the LLE website [46]. 22
- 3.5 An example of a directed graph generated by nearest neighbor associations. Here point 2 is a neighbor of point 1, point 3 is a neighbor of point 7, etc. In this case, the directed graph forms two strongly connected components: points 1, 2, 5, and points 3, 4, 6. In analyzing this data set, we would use LLE separately on each of these components and would remove point 7, which is not strongly connected to any other point. 23

- 4.1 Relationship of the sensed signal \tilde{h}_m to the mean soma potential \tilde{h}_e . In (a), an excitatory input synapses near the soma, causing a depolarization in \tilde{h}_e . Similarly, the dipole current generated by this synaptic event involves current sources near the cortical surface, which are manifested as a depolarization in the electrode measurement \tilde{h}_m . However, when the excitatory input occurs near the surface as in (b), the current dipole is reversed, causing opposite deflections in \tilde{h}_e and \tilde{h}_m . This figure is modeled after Box 46-1 in [18]. 27
- 4.2 Comparison of \tilde{h}_m (solid) and \tilde{h}_e (dashed) with two sets of parameters. At typical levels of excitation where $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$, the two signals are negatively correlated (top subfigure), as predicted by our physiological analysis. This negative correlation is especially noticeable if the synaptic time delay is taken into account by shifting \tilde{h}_e slightly to the left. However, when we simulate a strong excitatory input near the soma by setting $P_{ee} = 1000.0$, the signals become positively correlated (bottom subfigure). This was also predicted by the physiology of the cortex, and thus helps justify our model of \tilde{h}_m . Note that for both sets of parameters, the signal offsets (means) were removed to facilitate a direct comparison. 30
- 4.3 Comparison of the electrode measurement h_m (mV, solid) and the mean soma potential h_e (mV, dashed) at levels of subcortical excitation that cause seizure-like oscillations, $P_{ee} = 548.0$ and $\Gamma_e = 0.8 \times 10^{-3}$. Here we see that the two signals are negatively correlated, but have very similar dynamics. This indicates that it should be possible to perform feedback control using our new measurement model \tilde{h}_m (as we did previously with \tilde{h}_e in [22]), although a positive gain may be necessary. 31
- 4.4 Results for proportional control. (a) Plot of h_e [mV] in space and time. In the first 0.25 seconds, we see a seizure wave traveling on the simulated cortex because we have set the excitation parameters to $P_{ee} = 548.0$ and $\Gamma_e = 0.8 \times 10^{-3}$. At 0.25 seconds (dashed line), the proportional controller (4.4) is turned on, and we see that the seizure-like waves are immediately suppressed. For this simulation, we used \tilde{h}_m as the electrode measurement, and control was applied via five independent electrodes of width 11.2mm with a profile defined by the hyperbolic tangent function. The controller gains were $a_{\max} = 2$ and $b = 0$. (b) Plot of the control effort u [mV] associated with the simulation in (a). This shows the geometry of the five electrodes; in this case, they apply only negative signals. Note that, for clarity, the grayscale in (b) is the opposite of the one used in (a). 33
- 4.5 Results for a proportional controller with non-zero offset. (a) Plot of h_e [mV] in space and time. All parameters and settings are the same as in Figure 4.4, except $b = -0.3$. Again, the seizure-like waves are suppressed. In this case, the remaining fluctuations of h_e are larger than they were for proportional control (Figure 4.4), but now the controller is applying both positive and negative potentials. (b) Plot of applied voltage u [mV] for the five electrodes used in the simulation. 34

- 4.6 Results for differential control. (a) Plot of h_e [mV] in space and time. Excitation parameters and electrode geometry are the same as those used in Figure 4.4. At 0.25 seconds (dashed line), the differential controller (4.5) is turned on, and the seizure-like oscillations are suppressed. Again, the controller applies both positive and negative potentials; however, in this case, the frequency is higher than that of the seizure waves. The controller gains were $a_{\max} = 5$, $b = 0$, and $\tau_d = 20\text{ms}$. (b) Plot of applied effort u [mV] for the simulation of differential control. 35
- 4.7 Results for total integral feedback control. (a) Plot of h_e [mV] in space and time. Simulation parameters and electrode geometry are the same as those in Figure 4.4. The controller gains were $a_{\max} = 8$, $b = -0.1$, and $c = -8$. (b) Plot of applied effort u [mV] in space and time, corresponding to the simulation of feedback control in (a). The potential applied by the electrodes is zero until $t = 0.25$ seconds, when the controller is turned on (dashed line). Then each electrode oscillates between positive and negative signals, indicative of the “charge-balanced” nature of the control. For this simulation, the average magnitude of net applied effort over all five electrodes was 0.23mV. This value will approach zero as t increases, and it can also be reduced by increasing the integral control gain, c 36
- 4.8 Results for total integral feedback control at a single point in space. (a) Plot of h_e in time extracted from Figure 4.7(a) at 100.8mm. (b) Plot of control effort u in time extracted from Figure 4.7(b) at 100.8mm. Here we see that the control effort oscillates between positive and negative values in a charge-balanced manner. 37
- 4.9 Seizure waves for three values of the inverse corticocortical length scale λ_e : (a) 5.6, (b) 11.2, and (c) 22.4. An increase in λ_e corresponds to shorter, denser connections between neurons; this results in slower, higher amplitude waves over a shorter length of cortex. As with previous simulations, we used $\Gamma_e = 0.8 \times 10^{-4}$, $\alpha = 1.6$, and a gaussian distribution of P_{ee} with a maximum value of 548.0 at the center of the interval. 38
- 4.10 (a) Traveling waves in h_e for $\lambda_e = 11.2$, with no feedback control. This simulation used $\Gamma_e = 0.001$, $P_{ee} = 548.0$ and $\alpha = 5.0$. The waves travel from right to left with an approximate frequency of 9Hz and speed of 3m/s. (b) Traveling waves from (a), with the addition of one electrode performing proportional feedback control as described in (4.4). Here $a_{max} = 2$, $b = -0.1$, and the electrode width is 20.16 mm. 39

- 4.11 Change in standard deviation of h_e relative to uncontrolled seizure waves, for (a) $\lambda_e = 5.6$, (b) $\lambda_e = 11.2$ (typical value), and (c) $\lambda_e = 22.4$. After subtraction, the signals were smoothed with a 10-point moving average for ease of presentation. Each subfigure has traces for three different electrode sizes: 3.36mm, 10.08mm, and 20.16mm; the electrodes applied proportional control with $a_{\max} = 2$ and $b = -0.1$. Here we see that larger electrodes provide a greater reduction in standard deviation underneath the electrode. They also lead to a greater increase in standard deviation in front of the electrode and a greater decrease after the trailing edge. Note that the waves travel from right to left. 40
- 4.12 Close-up views of the results in Figure 4.11 showing the variation beneath the electrode. Note that the degree of control is greater at the trailing (left) edge of the electrode, as evidenced by the larger decrease in standard deviation. In addition, larger electrodes provide a greater overall decrease in standard deviation while using the same control gain. 41
- 4.13 (a) Change in standard deviation for three different values of λ_e , all with an electrode size of 3.36mm. As in previous figures, $\lambda_e = 5.6, 11.2, \text{ and } 22.4$. The data was taken directly from Figure 4.11. (b) A close-up view of the results in (a). These plots show that the spatial spread of the applied signal beyond the edges of the electrode is essentially independent of λ_e . The only exception to this is the trailing (left) edge of the electrode with $\lambda_e = 22.4$, which shows an overshoot in standard deviation before returning to uncontrolled levels. The pictures for other electrode sizes are qualitatively similar. 42
- 5.1 Examples of traditional EEG sleep stages, taken from PhysioBank data set sc4002e0: (a) awake, (b) Stage 1, (c) Stage 2, (d) Stage 3, (e) Stage 4, and (f) REM. Here, movement of the subject and EEG electrode appear to mask the typical awake behavior. As sleep deepens from stage 1 to stage 4, there is generally an increase in amplitude and a decrease in frequency. REM sleep, characterized by rapid eye movements, has a very low amplitude and resembles stage 1. 44
- 5.2 Scaled features of EEG data set sc4002e0, as described in Section 5.2.1. The subfigures show power in the (a) delta and (b) theta bands, (c) variance, (d) spindle score, (e) maximum height of the power spectrum in the alpha band after subtraction of a linear estimate, and (f) high power fraction. Subfigure (g) shows the hypnogram of the EEG data, where the number and color indicate the sleep stage: awake (0, black), stage 1 (1, yellow), stage 2 (2, green), stage 3 (3, cyan), stage 4 (4, blue), and REM (5, red). The features were calculated for the data from epochs 1597-1774 in 30-second windows with no overlap. 48

- 5.3 (a) Results of applying LLE to EEG data using the six features in Figure 5.2. The features were calculated for 30-second non-overlapping windows of data and the resulting 6-dimensional points were embedded in 2D space using LLE with $k = 13$; therefore, each point in this figure represents 30 seconds of EEG that has been characterized by the six features. The color and shape indicate sleep stage based on manual scoring: awake (black +), stage 1 (yellow \square), stage 2 (green \star), stage 3 (cyan \triangle), stage 4 (blue \ast), and REM (red \circ). (b) LLE output dimensions Y_1 and Y_2 versus time, for the results shown in (a). This demonstrates that LLE provides a low-dimensional output where the sleep state changes smoothly over time. 49
- 5.4 LLE results on sleep EEG data before (a) and after (b) removal of eight weakly connected points. The six features were power in the delta, theta, and gamma bands, total power, maximum height of PSD above a linear estimate in the alpha band, and low power fraction, and we used $k = 13$. As before, the color and symbol indicate sleep stage: awake (black +), stage 1 (yellow \square), stage 2 (green \star), stage 3 (cyan \triangle), stage 4 (blue \ast), and REM (red \circ). Note the dramatic improvement in separation between sleep stages when LLE is done on only one strongly connected component in (b). 50
- 5.5 Variation of five features as the surface of the sleep manifold is traversed in $L-\Delta h_e^{rest}$ space. Each feature has been scaled by its RMS value and depicted in grayscale, with white indicating the lowest values and black representing the highest values. (a) The steady state values of h_e from the sleep manifold in Figure 2.4. The black points represent the upper REM branch, the white points represent NREM, and the fold is located at roughly $L = 1.2$. The other subfigures show (b) power in the delta band, (c) permutation entropy, (d) maximum height of PSD above a linear estimate in the alpha band, (e) low power fraction, and (f) high power fraction. These five features use α as defined in (5.3). They show that the representation of REM and NREM in the model is consistent with the characteristics of sleep EEG. 52
- 5.6 Variation of the features from Figure 5.5 when they are applied to sleep EEG data, rather than model data. The sample of EEG data was taken from sc4002e0, and each feature has been scaled by its RMS value. The subfigures show (a) power in the delta band, (b) permutation entropy, (c) maximum height of PSD above a linear estimate in the alpha band, (d) low power fraction, (e) high power fraction, and (f) hypnogram of the EEG data. The colors and numbering for the hypnogram are the same as those used for Figure 5.2. Note that the values of these features (relative to sleep stage) are consistent with the model results in Figure 5.5. 53

5.7 (a) LLE results for a hybrid data set containing both sleep EEG data and numerical solutions of the cortical model. We used the five features from Figure 5.5 and set $k = 14$. The rings represent EEG data and are colored by sleep stage. While the analysis included 1200 windows of EEG data, only 500 are displayed here for clarity. The solid dots represent data from the model; they are colored based on the mean value of h_e at that point, where red represents the highest (REM) values, and dark blue marks the lowest (NREM) values. Note that the data and model points overlap in the output space and that the arrangement of sleep stages is very similar. (b) LLE results showing the EEG data *only*, using the same colors and symbols as Figure 5.4. This allows us to see that the data has been roughly separated by sleep stage. 54

5.8 Association between EEG data set sc4002e0 and the sleep manifold. Each picture shows the sleep manifold in $L-\Delta h_e^{rest}$ space, with a heavy black line to indicate the location of the fold. (a) Histograms of nearest neighbors for (i) waking, (ii) REM, (iii) stage 1, (iv) stage 2, (v) stage 3, and (vi) stage 4 sleep. The shading of each square indicates the number of times that location on the sleep manifold was a nearest neighbor of EEG data in that stage. For example, (vi) shows that stage 4 sleep most often associates itself with the lower NREM branch of solutions leading up to the fold. (b) A composite picture of the results in (a), where each location is colored based on the sleep stage with the most neighbors at that point, relative to the total number of neighbors associated with that stage. Again, we use stage 1 (yellow), stage 2 (green), stage 3 (cyan), stage 4 (blue), and REM (red). The intensity of color is scaled based on the percentage of neighbors that come from that stage; the more saturated the color, the greater the percentage. Waking points were excluded. 55

5.9 Composite plots for EEG data sets (a) sc4012e0, (b) sc4102e0, and (c) sc4112e0, when they are projected onto the LLE results from Figure 5.4(b), as described in Section 5.3.4. These pictures are analogous to Figure 5.8(b) and use the same color scheme. Note that the results are consistent with those for sc4002e0 in Figure 5.8(b); over various subjects, the sleep stages are generally associated with the same positions on the sleep manifold. 57

5.10 Combined association between all four EEG data sets and the sleep manifold. The data set sc4002e0 was directly compared to the cortical model using LLE, and the remaining three data sets were projected onto those results as described in Section 5.3.4. (a) Total histograms of nearest neighbors, separated by sleep stage; these were calculated by summing the histograms from all four EEG data sets. The pictures show awake, REM, and stages 1-4 in (i) through (vi), respectively. (b) The total composite picture for all four data sets. This was generated from the histogram data in (a) and is analogous to Figures 5.8(b) and 5.9. Again, this is consistent with previous results and shows the regions of the sleep manifold most closely associated with each sleep stage. 58

A.1 Plot of the same Brownian path in time at three different step sizes. 68

A.2 Plot of the same stochastic input over a strip of cortex at three different step sizes. 69

List of Tables

2.1	Dimensionless variables and parameters of the SPDE cortical model used for simulations of seizure waves. Values were taken from Kramer <i>et al.</i> [23]. For descriptions and values of the dimensional variables please refer to [55]. Note that L and Δh_e^{rest} are varied only for simulations of the sleep cycle; in studies of seizure behavior, they will always take the values 1 and 0, respectively. . .	14
2.2	Dimensionless variables and parameters of the SPDE cortical model used for simulation of the sleep cycle. The second column lists the definition in terms of the dimensional quantities in [61]. Values for the dimensional parameters were taken from that reference, with the exception of γ_{ia} which was chosen to be $90s^{-1}$	15
4.1	Values of the coefficients for equation (4.2). The first column of values was estimated as described here, while the second column was derived using probabilistic methods in [28].	29

Acknowledgments

First, thank you to the members of my dissertation committee: Andrew Szeri, Alexandre Bayen, Karl Hedrick, and Bruno Olshausen. Their guidance and support has been crucial to my success. Thanks also to my valued collaborators, James Sleigh at the Waikato Clinical School in New Zealand and Heidi Kirsch at the UCSF Epilepsy Center, for their patience and generosity.

There are several people I would like to acknowledge for their contributions to specific projects. Thank you to Bruce Gluckman and Steve Schiff for insightful discussions regarding electrode measurements. I would also like to recognize Kelly Clancy and Albert Kao for work which served as the starting point for chapter 5, done as part of a course in computational neuroscience at UC Berkeley taught by Professor Bruno Olshausen. In addition, Savas Tasoglu contributed to this project via his work on locally linear embedding.

Finally, thank you to all current and past inhabitants of 6112 Etcheverry Hall and my colleagues and friends in the Mechanical Engineering Department at UC Berkeley.

The work presented in this dissertation was supported through a National Science Foundation Graduate Research Fellowship. It was also supported, in part, by a Mary Elisabeth Rennie Epilepsy and Epilepsy-related Research Grant.

Chapter 1

Introduction

1.1 Seizures and epilepsy

1.1.1 Clinical presentation and diagnosis

Epilepsy is a neurological disorder that has been the subject of widespread research, but its varying etiologies and clinical presentations make it a challenging subject. A patient is given a diagnosis of epilepsy when he or she experiences recurrent, unprovoked seizures [18], where a seizure is broadly defined as a “brief, strong surge of electrical activity [that] affects part or all of the brain” [60]. The resulting symptoms and their severity vary widely from patient to patient and can often be mistaken for other conditions. Thus, epileptic seizures are often difficult to identify, even by experienced clinicians. When considering a diagnosis of epilepsy, doctors must account for a large number of factors, including the underlying cause of the abnormal electrical activity, age of onset, family history, and genetics [18].

In many cases, the cause is easily identifiable and the abnormalities are localized in one area of the brain. This is often true for seizures stemming from physical trauma, tumors, stroke, lead poisoning, infection, or problems with prenatal development [18]. However, in 70% of people with epilepsy, no cause can be found [60]. These cases are defined as *idiopathic* and are not accompanied by a local brain abnormality; in these situations, physicians must rely on clinical observation for the classification and diagnosis upon which treatment is based.

Based on these observations, seizures can be divided into two categories: partial and generalized. Partial seizures begin in a small group of neurons called a *focus*, and the symptoms experienced by the patient depend on the location of that focus, e.g. abnormal electrical activity in the motor area of the cortex can lead to twitching or jerking of the limbs. These seizures may or may not be accompanied by an alteration of consciousness. Shortly before partial seizures begin, the patient often experiences symptoms associated with an *aura*; for example, he or she may perceive a specific smell, have a sudden feeling of fear, or experience a sinking feeling in the pit of the stomach [18].

In contrast, generalized seizures involve both hemispheres of the brain. They are not preceded by an aura, and they do not begin at a localized focus. Generalized seizures can be either convulsive or nonconvulsive based on the physical movements (or lack thereof) during the seizure. Convulsive generalized seizures are perhaps most familiar to the general public; they last several minutes and are often characterized by a tonic phase (where the muscles

are rigid) followed by a clonic phase characterized by jerking of the extremities [18]. The most common example of a nonconvulsive generalized seizure is the absence seizure, often seen in children. These seizures are short, generally lasting less than 10 seconds, and cause a loss of consciousness while retaining muscle tone [18].

1.1.2 Seizure onset during sleep and waking

One of the most troubling aspects of generalized seizures is their unpredictability; in addition to the physical trauma, this leads to social implications including the impairment of a person's ability to drive or work. Stress, lack of sleep, or flashing lights at certain intensities and frequencies can sometimes increase the likelihood of seizures, but many epileptic patients are unable to identify a specific trigger [60].

Patterns of seizure onset have also been found in relation to the sleep/wake cycle, although this connection has yet to be fully understood. The analysis of human sleep generally involves categorization of EEG measurements into five groups: rapid eye movement (REM) sleep and stages 1-4 (also called non-REM stages) [36]. The probability of seizure occurrence during sleep appears to depend not only on sleep stage, but also on the type of seizure, the location of seizure onset, and the presence of arousals or interictal epileptiform discharges (IEDs). For example, some epileptic syndromes have a tendency to produce seizures during sleep (e.g. frontal lobe epilepsy), while others are associated with seizures during waking (e.g. absence epilepsy) [44, 8]. Seizures during sleep occur primarily in non-REM (NREM) sleep stages, while rapid eye movement (REM) sleep inhibits epileptic activity [44, 16, 8]. Within NREM sleep, IEDs are prevalent during deep sleep (stages 3 and 4), whereas most seizures arise out of the lighter stages 1 and 2 [32, 44]. It has been hypothesized that the transitions between these sleep stages are critical points in seizure generation, especially during the descent to deep sleep after an arousal [32]. However, the knowledge of these patterns has not yet led to the ability to anticipate the onset of a seizure during sleep.

The standard method of dividing sleep stages into five *discrete* groups adds to the difficulty of quantifying the precise relationship between seizure onset and sleep state. For example, if a seizure occurs during stage 2 sleep, exactly how long was the subject in stage 2 before seizure onset? Was the subject descending to deeper stages of sleep or arising from them? Was a transition imminent when the seizure occurred?

The use of a mathematical model of the human sleep cycle may allow such questions to be answered by providing a *continuous* spectrum of sleep states, ranging from REM to the deepest slow-wave sleep. If the model can be directly associated with human sleep EEG data, it would be possible to track the subject's state to identify the stage as well as changes in sleep depth and proximity to transitions. Ideally, this would be done in real-time, where the state (and likelihood of seizure onset) is continuously determined as the subject sleeps. Assuming this process is consistent over various subjects and robust to non-standard sleep cycles and periods of waking, it could represent a significant step toward seizure prediction and suppression during sleep.

1.1.3 Treatment

Due to the unpredictable nature of seizures, the best treatments are often preventative. The most common method is medication; it is estimated that over 50% of epileptic patients are able to completely control their seizures with antiepileptic drugs, and another 20-30% see a reduction in the number of seizures [60]. The remaining 20% have *intractable* epilepsy that does not respond to medication. These patients often seek out alternative treatments, which tend to be much more invasive. For example, he or she may elect to have surgery to resect the seizing portion of the brain. Success for this type of procedure, defined as “no seizures with loss of consciousness,” is reported in 30-70% of cases [60]. However, it can only be performed when the seizures occur in a part of the brain unrelated to basic functions such as speech, memory, or eyesight. Another treatment option for partial seizures is vagus nerve stimulation. This involves the implantation of a small device to deliver periodic electrical pulses to the brain via the vagus nerve in the neck; the stimulation is generally delivered in an open-loop manner, but it is sometimes possible for the patient to activate the device with a magnet if they feel a seizure coming on. This type of treatment is unlikely to stop seizures completely, but roughly one third of patients will see a great improvement in seizure control [60].

Cortical stimulation is a new form of treatment that is currently the subject of much research. It utilizes subdural cortical surface electrodes and/or depth electrodes to detect seizures and then to apply an electrical signal to disrupt the abnormal electrocorticogram (ECoG) activity. As opposed to vagal nerve stimulation or medication, the electrical signal is directed locally to the site of the seizure and can be applied as needed with great temporal specificity [57]. Most importantly, it does not require removal of any part of the brain. With this treatment method, stimulation parameters can be tailored to the needs of each patient, and the device can be switched off very quickly if adverse side effects occur [57]. The stimulation can be applied in an open-loop manner, where electrical pulses are delivered at predetermined intervals with set intensities, or in a closed loop manner, where the stimulus is calculated based on continuous measurements of the brain’s electrical activity. A third protocol, termed “responsive cortical stimulation,” is a hybrid of these two methods. It calls for stimulation only when a seizure has been detected (in this sense, it is a closed loop system), but the applied signals are predetermined (as they are for open loop stimulation) [57].

A limited set of related experiments has been done on humans. The first use of cortical stimulation on a human subject was published in 1954; it was reported that an applied stimulus could locally depress electrical activity and, in some cases, cause a complete flattening of the ECoG signal in distant cortical regions [40]. This was also seen later by physicians performing resective surgery for epilepsy. In an attempt to map regions of the cortex used for critical functions such as speech and memory, surgeons apply a small electrical signal to different areas of the cortex and observe the response. If the amplitude of the applied signal is large enough, an *after-discharge* will be produced. This electrical phenomenon mimics the spontaneous epileptiform activity at the beginning of a focal seizure and may evolve into a full-blown seizure [57]. It was empirically found that additional electrical stimulation could terminate these after-discharges once they had begun, and studies were performed to quantify the characteristics of the signal needed to do this [25, 33]. Others used interictal

epileptiform discharges as a measure of seizure activity; they studied the reduction in IEDs due to applied electrical signals of varying frequencies [20]. Most recently, clinical trials were initiated on a fully integrated implantable device that uses responsive cortical stimulation for seizure suppression. [57].

To our knowledge, closed loop control has not yet been implemented in humans, but it is currently being studied in animal models of epilepsy. It has been shown that the application of electric fields to rat cortex *in vitro* can modulate the behavior of seizure-like waves [42]. *In vivo* experiments on rats demonstrated that stimulation via proportional feedback can temporarily suppress seizure activity [13]. A subsequent set of experiments showed that an increase in the amplitude of the proportional control feedback gain corresponds to decreases in both seizure amplitude (measured as a reduction in the amplitude variance) and Teager energy [5]. The Teager energy can be reduced by a decrease in the amplitude of a signal or a lowering of the frequencies in its power spectrum.

1.2 Computational Neuroscience, Epilepsy, and Sleep

While animal experiments offer greater freedom over human studies, mathematical models provide the most flexibility of all. They present an opportunity to study many different aspects of epilepsy, from the contributions of a single type of ion channel to the combined oscillatory behavior of a thousand networked neurons. The type of model must be chosen to suit the research goal. For example, small-world networks of simple integrate-and-fire neurons have been shown to exhibit self-sustained oscillations similar to a seizure [35]. While this may provide insight into the network topologies that facilitate seizure behavior, it does not contain enough details to study the physiological generation of the seizure. On the other end of the spectrum, the NEURON simulation package has the ability to model neurons in extreme detail based on experimental data. Researchers have used this to study the actions of anti-convulsant medications at the single cell level, the effect of specific ion channels on neuronal excitability and bursting, and much more [4].

One type of model that falls in between these two extremes is the mean-field model. Its variables, representing quantities that are averaged over the millimeter scale, are comparable to the mesoscale measurements of the electroencephalogram (EEG) or ECoG electrodes. Therefore, this mathematical representation of cortical tissue is well-suited to the study of brain states described by EEG signals, including seizures, anesthesia, and sleep. However, because the contributions of individual neurons are lost due to the assumption of spatial averaging, these models are not useful for studying seizure-related mechanisms at the cellular level (or below). They can be used to simulate the action of many thousands of neurons, but unlike the small-world network example above, individual synapses and connection strengths are not represented. Instead, the model characterizes the connection between an averaged excitatory population and an averaged inhibitory population. While this rules out the possibility of studying network topologies, it provides a great advantage in terms of efficiency. A mean-field model has mathematical equations relating to a neuronal *population*, whereas a large-scale network model must use several equations for *each neuron*. Thus, mean-field models are orders of magnitude smaller in terms of the number of differential equations that must be solved.

Mathematically, the cortical model used in this manuscript is a 12th order set of stochastic partial differential equations (SPDEs) with one spatial dimension. It can support seizure-like oscillations that are qualitatively and quantitatively similar in frequency of maximum power and propagation speed to those seen in humans with epilepsy [21]. Also, by adding two parameters to represent neurotransmitters important for the regulation of sleep, this cortical model contains a mathematical representation of the human sleep cycle. In what follows, we take advantage of these two properties by using the mean-field, mesoscale model of the cortex to study both seizure control and sleep dynamics.

1.2.1 Feedback control of epileptic seizures

The experiments on closed loop feedback control of epileptic seizures in animals have provided promising results, but they could benefit from the application of a mathematical model, such as the one used here. We previously demonstrated that various methods of closed loop feedback control can suppress simulated seizures [22], and we added the capability of looking at spatial properties of feedback control, such as electrode size and spacing [29].

Our goal in the present work is to make this model of feedback control more biologically relevant. This will facilitate future comparison with experimental data. There are four key improvements in our approach:

1. We verify the strong convergence of the numerical solution to the SPDE model.
2. Based on the studies of convergence, we utilize a smaller step size in our simulations, thereby allowing the inclusion of a more detailed electrode profile.
3. We develop a better motivated model of the signal measured by an electrode on the cortical surface. This model is used to calculate the applied electrical signal for feedback control.
4. Feedback control is performed with a new algorithm incorporating an integral component. This ensures that the applied signal is charge-balanced, which is thought to minimize damage of cortical tissue.

Then, using this new formulation of feedback control, we study the relationship between the cortical length scale in the model and the size of the electrode needed to effectively suppress a seizure. This allows us to make recommendations for electrode design in experimental implementations of feedback control.

1.2.2 A continuous mapping of human sleep states

As was previously discussed, the relationship between sleep and epileptic seizures is complex and not fully understood. The description of sleep stages in a continuous manner may aid in the understanding of this subject. Here we are able to find a continuous mapping of EEG sleep states by direct comparison to the model of the sleep cycle contained within the mesoscale cortical model. Basic descriptions of NREM and REM sleep were previously identified in the model, and these sleep states were qualitatively compared to EEG data [52]. However, we aim to quantify this relationship.

First, we introduce a method of nonlinear dimensionality reduction called locally linear embedding, and we show that it provides the ability to distinguish between sleep stages when applied to EEG data. These results demonstrate reliable separation between REM and NREM sleep data and provide a smooth temporal progression through the various stages of sleep. We also present the concept of strongly connected components as a method of outlier rejection for EEG data and introduce a method for automatic selection of LLE parameters. Then, by performing LLE on a hybrid data set containing both sleep EEG and signals generated from the mathematical model, we are able to integrate the EEG and the model. This allows us to take any sample of sleep EEG data and determine its position within the continuous range of sleep states provided by the model. We show that this method provides consistent results for various subjects over a full night of sleep, and it could be done in real time as the subject sleeps.

1.3 Summary of Original Contributions

Using a mean-field, mesoscale model of the human cortex, we are able to make novel contributions relating to both feedback control of seizures and the continuous mapping of sleep states to aid in seizure prediction.

The simulations of feedback control presented here are an improvement over previous work in several ways. First, we provide recommendations for the maximum simulation step sizes based on mathematical analysis of convergence, and we include the capability to define a detailed electrode profile. Next, we introduce a new model for the measurement of a cortical surface electrode based on extracellular currents in the tissue. All previous work was based on the simplifying assumption that the intracellular soma potential of the neurons would be reflected in the sensor measurement. Lastly, we present a new control law that guarantees seizure suppression with a *charge-balanced* signal. This is of great importance for implementation of feedback control because this type of signal is thought to cause the least amount of damage when applied to cortical tissue.

The continuous mapping of EEG sleep states represents the combination of three components (LLE, sleep EEG data, and a mesoscale model of the sleep cycle) in a way that has never been seen before. To our knowledge, this is the first time LLE has been applied to sleep EEG data, the first time LLE has been used as a bridge between experimental data and a mathematical model, and the first time EEG data has been quantitatively compared to this mean-field model of the sleep cycle. We show that LLE has the ability to roughly separate sleep EEG by stage, and we demonstrate that the resulting embedding changes smoothly over time. Then, by jointly analyzing EEG data and model data using LLE, we quantify the mapping between these two data types. The results are consistent over multiple subjects with varying sleep cycles, and the analysis can be implemented in real-time. Because it results in a more detailed description of sleep states, this work may have implications for seizure detection and suppression during sleep.

Chapter 2

Mesoscale Cortical Model

2.1 Description of the model

To model the electrical activity of the human cortex, we choose a set of stochastic PDEs that has been developed and adapted over the past decade [27, 26, 54, 55]. The mesoscale nature of this model makes it well-suited to EEG-based applications such as seizure waves associated with epilepsy [21], sleep dynamics [61], and cortical transitions due to the application of anesthesia [53] because it is based on length scales similar to commercial electrode arrays. It is a *mean-field* model, implying that all of its variables represent spatially averaged properties of populations of neurons. This is similar to the manner in which an electrode provides a measurement based on the collective behavior of many neurons. We choose to implement the model in only one spatial dimension because this will be sufficient to capture the plane waves associated with epileptic seizures. A two-dimensional model would be needed to simulate more complex wave patterns (e.g. spiral waves), which have been observed in laboratory preparations [49].

In 2006, the equations were restated in a dimensionless form by Kramer *et al.* [23]. This is the formulation of the model we will use here; it is a system of eight coupled nonlinear

PDEs with stochastic inputs:

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

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

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

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

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

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

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

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

All variables are dimensionless and are functions of time (\tilde{t}) and one spatial dimension (\tilde{x}). The \tilde{h} state variable is the mean soma potential for a neuronal population, while \tilde{I} represents postsynaptic activation due to local, long-range, and subcortical inputs, and $\tilde{\phi}$ is a long-range (corticocortical) input. The subscripts e and i denote the excitatory and inhibitory neuron populations, respectively; variables with two subscripts represent the transmission of information from one population to another, e.g. \tilde{I}_{ie} is the postsynaptic activation of the excitatory population due to inputs from the inhibitory population. For descriptions of all model variables and parameters, please refer to Table 2.1 (for parameters to produce seizure waves) or Table 2.2 (for parameters to model the sleep cycle).

Note that two small changes have been made to the model in comparison to [23]. In equation (2.1) we have added the variable u to represent the signal applied by a cortical surface electrode for feedback control. This will be discussed further in Section 4.1.3. Also, in equations (2.1) and (2.2) the parameters L and Δh_e^{rest} have been added to mimic neuromodulators present during human sleep [52]. Changes in these parameters will allow us to model the sleep cycle, as discussed in Section 2.2.2; for simulations of seizure behavior, we will always choose $L = 1$ and $\Delta h_e^{rest} = 0$.

To appreciate the model as a whole, let us first look at the equations governing the excitatory neuron population, depicted graphically in Figure 2.1. Equation (2.1) for \tilde{h}_e is reminiscent of the leaky integrate-and-fire model of a neuron, where the derivative of the membrane potential equals the resting potential minus the membrane potential plus any existing current inputs [7]. Here, the resting potential is “1” due to the dimensionless nature of the system. The inputs are \tilde{I}_{ee} and \tilde{I}_{ie} , which evolve according to (2.3) and (2.5), respectively, based on three types of synaptic input: local, long-range, and subcortical.

- **Local inputs**, such as those from within the same macrocolumn, are represented by

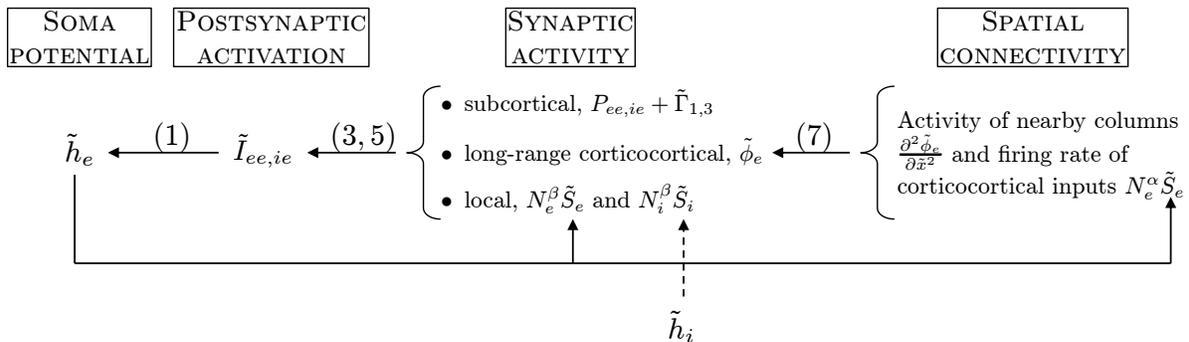


Figure 2.1: Flow chart representation of model equations (2.1), (2.3), (2.5), and (2.7), which govern the excitatory population. The boxes describe the physiological significance of the model variables and parameters listed beneath them. Note that feedback occurs through the sigmoid function \tilde{S}_e , which is a function of \tilde{h}_e . This population is also coupled to the inhibitory population through local inputs described by $\tilde{S}_i[\tilde{h}_i]$. For a cell-based depiction of the model, see [54].

terms of the form $N_e^\beta \tilde{S}_e[\tilde{h}_e]$, where \tilde{S}_e is a dimensionless sigmoid function:

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

This associates the mean soma potential of the excitatory population with its mean firing rate.

- **Long-range inputs** represent signals from other cortical macrocolumns and are defined by $\tilde{\phi}_e$. The behavior of this variable is governed by (2.7). Note the similarity of this equation to the standard PDE wave equation. It has been experimentally verified that cortical tissue can support wave propagation [49]; this idea is central to our simulation of epileptic seizures.
- **Subcortical inputs** are predominantly from the thalamus and contain both constant (P_{ee}) and stochastic ($\tilde{\Gamma}_1$) parameters. We define the stochastic term by

$$\tilde{\Gamma}_1 = \alpha_{ee} \sqrt{P_{ee}} \xi_1[\tilde{x}, \tilde{t}], \quad (2.10)$$

where α_{ee} is a constant and ξ_1 is zero mean, Gaussian white noise in time and one spatial dimension. When the SPDEs are solved numerically, the cumulative effect of this stochastic process will be Brownian motion. To ensure that the properties of this signal remain constant regardless of step size, we scale the discrete randomly generated numbers $R(m, n)$ by the simulation time step:

$$\xi_1[\tilde{x}, \tilde{t}] = \frac{R(m, n)}{\sqrt{\Delta \tilde{t}}}. \quad (2.11)$$

The variables m and n are indices of space and time, so a single point is represented by spatial position $\tilde{x} = m\Delta\tilde{x}$ at time $\tilde{t} = n\Delta\tilde{t}$. Note that the discrete update form of (2.11) is

$$\xi_1[\tilde{x}, \tilde{t}]\Delta\tilde{t} = \sqrt{\Delta\tilde{t}}R(m, n), \quad (2.12)$$

which will be used in all numerical experiments.

Thus, equations (2.1), (2.3), (2.5), and (2.7) govern the excitatory population, while the remaining equations represent the inhibitory population and have exactly the same form. Together they compose the full cortical model.

2.2 Representations of cortical states in the model

2.2.1 Seizure waves

The full set of parameters used to study seizure-like behavior in the model are listed in Table 2.1, but several are especially relevant to the following numerical studies. The parameter P_{ee} represents input from the population of subcortical excitatory neurons (such as those in the thalamus), and Γ_e denotes the influence of synaptic input on the mean soma potential. Changes in these parameters allow for transition between normal cortical function and the hyperexcited “seizure” state of the SPDE model. At low levels of excitation, corresponding to low levels of P_{ee} , the mean soma potential of the excitatory neurons \tilde{h}_e produces low amplitude random fluctuations similar to those seen in an ECoG measurement on the cortical surface. An example of this is shown in Figure 2.2 for a one-dimensional strip of cortex evolving in time; here we used $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$, and the value of h_e [mV] is represented by grayscale.

On the other hand, at increased levels of subcortical excitation, the simulated cortex develops large amplitude seizure-like oscillations. This is depicted in Figure 2.3, where the behavior is again represented by the mean soma potential of the excitatory population h_e as it varies in space and time. The seizure waves occur due to our choice of $\Gamma_e = 0.8 \times 10^{-3}$ and a Gaussian P_{ee} distribution with a maximum value of 548.0. They spontaneously arise in locations with sufficiently high P_{ee} (here, this “hot spot” is at $x = 100\text{mm}$) and travel outward until the level of excitation is too low to support them. This is why the waves terminate before they reach the edges of the simulation space.

These large-amplitude oscillations were found to be quantitatively and qualitatively similar to seizure waves seen in epileptic subjects, with respect to the frequency of peak power and wave propagation speed [21]. It was also shown that an increase in P_{ee} and decrease in Γ_e is not the only route to seizure activity in the model; these oscillations can develop through a variety of other parameter changes [23]. Our goal in Chapter 4 will be to suppress this pathological behavior via feedback control consisting of measurements from the cortical surface and the application of a potential based on those measurements. The variable u will represent this control input.

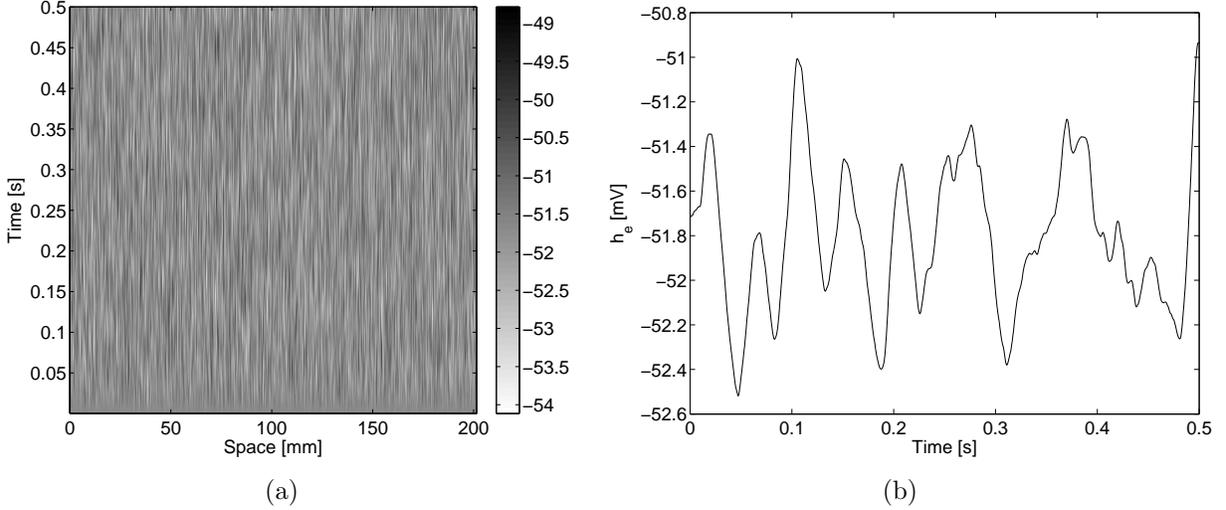


Figure 2.2: (a) Typical behavior of the cortical model, with parameters $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$. Here we simulate a 200mm strip of cortex over a period of 0.5 seconds, and the value of h_e is represented by grayscale. For these parameters, the model exhibits low amplitude stochastic behavior. (b) Plot of h_e versus time extracted from Figure 2.2(a) at 100.8mm.

2.2.2 Sleep states

To simulate sleep states, we again adopt the dimensionless formulation of the model as described in [23]; however, here we use parameters Δh_e^{rest} and L to equations (2.1) and (2.2) represent neuromodulators that regulate the natural sleep cycle, as was done in [52]. More specifically, they mimic the actions of adenosine and acetylcholine (ACh). First, adenosine acts to reduce the resting potential of excitatory cells, thus making them less likely to fire; ACh does the opposite by raising the resting potential. These actions are represented in the model by Δh_e^{rest} , which adds directly to the resting potential of the excitatory population (disguised as a “1” in the dimensionless equations). Second, ACh decreases the amplitude of the excitatory postsynaptic potential, effectively reducing the synaptic gain. In the model, this corresponds to a reduction in the effect of synaptic currents \tilde{I}_{ee} and \tilde{I}_{ei} ; therefore, the parameter L is multiplied by these quantities to simulate a change in synaptic gain. Lastly, as was done in [52], we take the mean excitatory soma potential \tilde{h}_e to be representative of cortical activity; we will compare this variable to EEG measurements in the analysis of Chapter 5.

In order to visualize the model of the sleep cycle based on these new parameters, we look at steady-state solutions of h_e (without stochastic input) as L and Δh_e^{rest} are varied. These solutions create what we will refer to as the “sleep manifold” (Figure 2.4). Notice that, for most parameter values, there is only one steady state solution. However, in certain cases, there are three solutions (two stable and one unstable), causing the manifold to fold over on itself. This fold is seen on the left side of Figure 2.4. In this model, the top branch of solutions on the manifold is intended to be representative of REM sleep. Starting at this point, we can imagine that during sleep, there is a gradual descent to deep slow-wave sleep

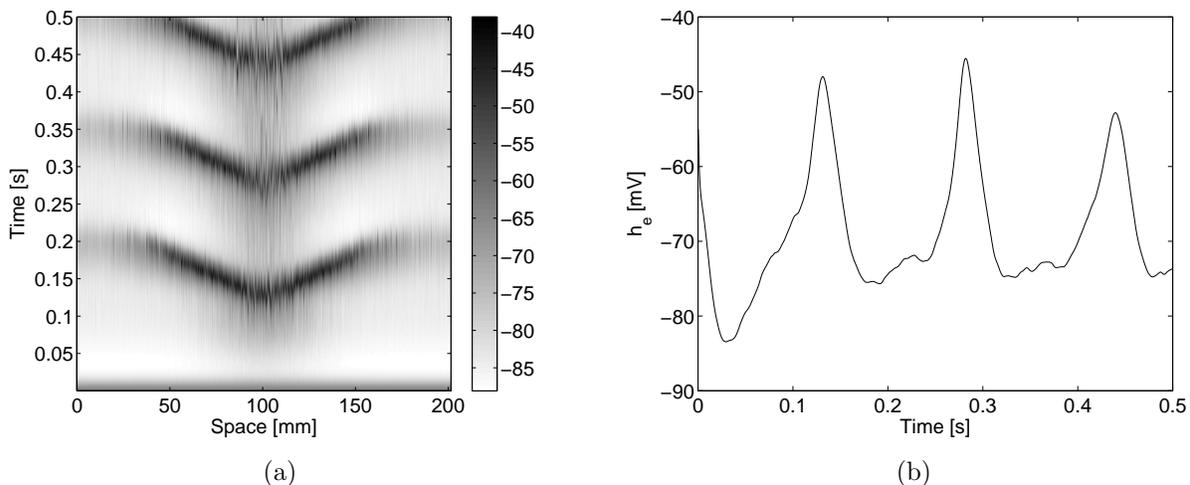


Figure 2.3: (a) Seizure waves traveling on the simulated cortex, with parameters $P_{ee} = 548.0$ and $\Gamma_e = 0.8 \times 10^{-3}$. The characteristics of the wave are determined by the distribution of P_{ee} ; here it is a Gaussian curve, so the wave starts in the center where P_{ee} is at its maximum (548.0) and propagates outward until the level of excitation is too low to support it. In this case, no feedback control is applied, so the waves will reoccur indefinitely. (b) Plot of h_e in time extracted from Figure 2.3(a) at 100.8mm.

by following a trajectory down the right side of the manifold where there is only one steady state solution. This happens in a smooth continuous manner. Then the quick transition from slow-wave sleep to REM is simulated by a jump across the fold from the bottom branch of solutions to the top branch. This mimics the rapid transition from deep sleep to REM that is observed in human EEG recordings. This process of gradually moving from REM to NREM sleep and then quickly jumping back to REM represents one sleep cycle in the model.

This model has been previously studied. Steyn-Ross *et al.* calculated the EEG total power, fractions of high and low power, and correlation time exhibited by the model at the transition from slow-wave sleep to REM; it was found that they qualitatively matched both human clinical sleep recordings and cortical measurements from a cat [52]. The model was also studied in two spatial dimensions to investigate stable oscillatory states similar to slow-wave sleep, and it was shown that a transition from one state to another can occur due to stochastic fluctuations [62]. Lastly, Wilson *et al.* interpreted the k-complex as a transient shift from a stable low-firing state to an unstable high-firing state and used this model to demonstrate the mechanism by which the transition may occur [61]. In Chapter 5, we will quantitatively compare this model to sleep EEG data; for any given sample of EEG, our aim will be to find its associated position on the sleep manifold.

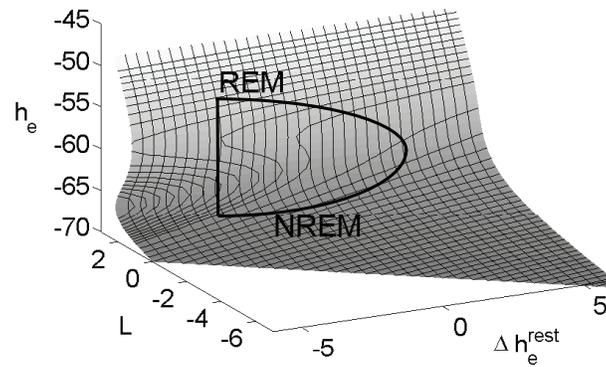


Figure 2.4: The manifold of steady states in h_e from the mesoscale cortical model, hereafter referred to as the “sleep manifold.” The parameters L and Δh_e^{rest} represent the actions of adenosine and acetylcholine, neuromodulators that vary over the course of the human sleep cycle. The manifold has two stable solutions on its left side; a jump from the bottom solution to the top solution represents the fast transition between NREM and REM sleep. However, the slow transition from REM to NREM occurs smoothly down the right side of the manifold, where there is only one set of solutions.

Symbol	Definition	Description	Typical value
$\tilde{h}_{e,i}$	$h_{e,i}/h^{rest}$	Spatially averaged soma potential for neuron populations	–
$\tilde{I}_{ee,ei}$	$\frac{I_{ee,ei}\gamma_e}{G_e \exp(1)S^{\max}}$	Postsynaptic activation due to excitatory inputs	–
$\tilde{I}_{ie,ii}$	$\frac{I_{ie,ii}\gamma_i}{G_i \exp(1)S^{\max}}$	Postsynaptic activation due to inhibitory inputs	–
$\tilde{\phi}_{e,i}$	$\phi_{e,i}/S^{\max}$	Long-range (corticocortical) input to e and i populations	–
\tilde{t}	t/τ	Time (dimensionless)	–
\tilde{x}	$x/(\tau\tilde{v})$	Space (dimensionless)	–
$\Gamma_{e,i}$	$\frac{G_{e,i} \exp(1)S^{\max}}{\gamma_{e,i} h_{e,i}^{rev} - h^{rest} }$	Influence of synaptic input on mean soma potential	1.42×10^{-3} , 0.0774
$h_{e,i}^0$	$h_{e,i}^{rev}/h^{rest}$	Reversal potential	–0.643, 1.29
$T_{e,i}$	$\tau\gamma_{e,i}$	Neurotransmitter rate constant	12.0, 2.6
$\lambda_{e,i}$	$\tau\tilde{v}\Lambda_{ee,ei}$	Inverse length scale for corticocortical connections	11.2, 18.2
$P_{ee,ei}$	$p_{ee,ei}/S^{\max}$	Subcortical input from excitatory population	11.0, 16.0
$P_{ie,ii}$	$p_{ie,ii}/S^{\max}$	Subcortical input from inhibitory population	16.0, 11.0
$N_{e,i}^\alpha$	–	Number of distant (corticocortical) connections from excitatory populations to e and i populations	4000, 2000
$N_{e,i}^\beta$	–	Number of local synaptic connections from e and i populations	3034, 536
$\tilde{g}_{e,i}$	$g_{e,i}h^{rest}$	Slope at inflection point of sigmoid function \tilde{S}_e	–19.6, –9.8
$\tilde{\theta}_{e,i}$	$\theta_{e,i}/h^{rest}$	Inflection point for sigmoid function \tilde{S}_e	0.857, 0.857
L	–	Synaptic gain due to the action of Ach during sleep	1
Δh_e^{rest}	–	Change in resting potential during sleep due to adenosine and Ach	0

Table 2.1: Dimensionless variables and parameters of the SPDE cortical model used for simulations of seizure waves. Values were taken from Kramer *et al.* [23]. For descriptions and values of the dimensional variables please refer to [55]. Note that L and Δh_e^{rest} are varied only for simulations of the sleep cycle; in studies of seizure behavior, they will always take the values 1 and 0, respectively.

Symbol	Definition	Description	Typical value
$\tilde{h}_{e,i}$	$V_{e,i}/V_{e,i}^{rest}$	Spatially averaged soma potential for neuron populations	–
$\tilde{I}_{ee,ei}$	$\Phi_{ee,ei}/Q_e^{max}$	Postsynaptic activation due to excitatory inputs	–
$\tilde{I}_{ie,ii}$	$\Phi_{ie,ii}/Q_i^{max}$	Postsynaptic activation due to inhibitory inputs	–
$\tilde{\phi}_{e,i}$	$N_{ee,ei}^\alpha \Phi_{ee,ei}/Q_{e,i}^{max}$	Long-range (corticocortical) input to e and i populations	–
\tilde{t}	$t/\tau_{e,i}$	Time (dimensionless)	–
\tilde{x}	$x/(\tau_{e,i} v)$	Space (dimensionless)	–
$\Gamma_{e,i}$	$\frac{\exp(1)Q_{e,i}^{max}}{\gamma_{ea,ia}(V_{e,i}^{rev} - V_{e,i}^{rest})}$	Influence of synaptic input on mean soma potential	4.6875×10^{-4} , 0.0105
$h_{e,i}^0$	$V_{e,i}^{rev}/V_{e,i}^{rest}$	Reversal potential	0, 1.0938
$T_{e,i}$	$\tau_{e,i} \gamma_{ea,ia}$	Neurotransmitter rate constant	12.0, 3.6
$\lambda_{e,i}$	$\tau_{e,i} v \Lambda_{ea}$	Inverse length scale for corticocortical connections	11.2, 11.2
$P_{ee,ei}$	$\langle \phi_{ea}^{sc} \rangle / Q_e^{max}$	Subcortical input from excitatory population	25.0, 25.0
$P_{ie,ii}$	$\langle \phi_{ia}^{sc} \rangle / Q_i^{max}$	Subcortical input from inhibitory population	25.0, 25.0
$N_{e,i}^\alpha$	$N_{ee,ei}^\alpha$	Number of distant (corticocortical) connections from excitatory populations to e and i populations	3710, 3710
$N_{e,i}^\beta$	$N_{ea,ia}^\beta$	Number of local synaptic connections from e and i populations	410, 800
$\tilde{g}_{e,i}$	$\pi V_{e,i}^{rest} / (\sqrt{3} \sigma_{e,i})$	Slope at inflection point of sigmoid function \tilde{S}_e	–29.021, –19.347
$\tilde{\theta}_{e,i}$	$\theta_{e,i}/V_{e,i}^{rest}$	Inflection point for sigmoid function \tilde{S}_e	0.91406, 0.91406

Table 2.2: Dimensionless variables and parameters of the SPDE cortical model used for simulation of the sleep cycle. The second column lists the definition in terms of the dimensional quantities in [61]. Values for the dimensional parameters were taken from that reference, with the exception of γ_{ia} which was chosen to be $90s^{-1}$.

Chapter 3

Mathematical methods

This chapter provides an introduction to several techniques that will be used later in the manuscript. Tests of strong convergence provide step size recommendations that will be incorporated into the simulations in chapter 4. The statistical measurements described below will be used in Section 5.2 as a means of analyzing sleep EEG data. Then, later in chapter 5, locally linear embedding and the concept of strongly connected components will be used to map the EEG data to a mathematical model of the sleep cycle.

3.1 Strong convergence of numerical solutions

Before performing simulations of feedback control, we must ensure that we can obtain accurate numerical solutions to this system of SPDEs. We will use a predictor-corrector algorithm written in MATLAB, so the accuracy of the solution will be determined by our choice of step sizes in space and time. In addition to considering the system of equations and solution method, we shall assume that a typical cortical surface electrode is of order 1cm in diameter. While the previously used step size of 7mm [29, 22] may have accurately solved the differential equations, it was not small enough for sufficient spatial resolution of the behavior of the cortical tissue underneath the electrodes. We will use much smaller step sizes in order to achieve both of these objectives.

3.1.1 Method

To determine the magnitude of these step sizes, we will examine the strong convergence of solutions to the cortical model. This will be accomplished by generating *equivalent* Brownian paths at several step sizes and demonstrating that the solutions converge as the step size decreases (note that this differs from weak convergence, which looks at the expected value of the solution over all possible Brownian paths) [15]. This task will be complicated by the fact that both the stochastic inputs and the numerical solutions vary in space and time.

Recall from (2.12) that the grid of stochastic inputs is defined by R , which consists of M independent Brownian paths, each of length N . This corresponds to M points in space at a reference step size of $\Delta x = \Delta x_0$ and N points in time at a reference step size of $\Delta t = \Delta t_0$. Therefore, we denote individual points by $R(m, n)$, $\forall m = 1, \dots, M$ and $\forall n = 1, \dots, N$. Then

we can represent an *equivalent* Brownian path at a coarser step size $2\Delta t_0$ by adding together every two adjacent elements in time:

$$\tilde{R}(m, n) = R(m, 2n - 1) + R(m, 2n) , \quad (3.1)$$

where $n = 1, \dots, N/2$ [10]. We do not need any special scaling factors here because this combination of neighboring terms is consistent with the definition of a Brownian path. Similarly, we can represent the stochastic input at step size $2\Delta x_0$ by adding together adjacent elements in space and scaling to keep the variance constant [11]:

$$\tilde{R}(m, n) = \frac{1}{\sqrt{2}} (R(2m - 1, n) + R(2m, n)) , \quad (3.2)$$

where $m = 1, \dots, M/2$. The factor $1/\sqrt{2}$ is necessary because the stochastic inputs are independent in the spatial direction; it will be used for the *relative* scaling of the inputs at different step sizes for the purpose of determining convergence, but will not be present in a typical simulation of feedback control.

We can do a simple simulation in Matlab to verify that equations (3.1) and (3.2) allow us to generate equivalent Brownian paths at relative step sizes. Code for this purpose is provided in Appendix A.1, along with plots of the results. For a mathematical proof that equation (3.2) preserves the variance of the stochastic signal, please see Appendix A.2. Similarly, a Matlab simulation to illustrate the generation of Brownian inputs at *absolute* step sizes is given in Appendix B.

We are now in a position to directly compare numerical solutions at decreasing step sizes (e.g. $\Delta x = 4\Delta x_0, 2\Delta x_0, \Delta x_0$) under equivalent stochastic inputs. We want to verify that the solution converges as we approach Δx_0 .

3.1.2 Results

First, we look at the convergence in time using the method described above. We remove the spatial terms from the cortical model to reduce it to an ODE and then perform simulations with decreasing values of the time step. These indicate that the solution converges around $\Delta t = 5 \times 10^{-4}$ s (Figure 3.1). The two smallest time steps in the figure, $\Delta t = 5 \times 10^{-4}$ s and 2.5×10^{-4} s give very similar results for h_e .

Next, we study the convergence in space. We begin with a Δx that is smaller than 1cm because we desire to resolve the solution across an electrode. As Δx decreases, we see more detail present in the solution; however, this does not give us a clear indication of which step size to choose. The amount of improvement seems to be the same for each reduction in Δx . We solve this problem by looking at the numerical worst-case scenario – a sharp transition between uncontrolled cortex and a single electrode with proportional feedback. We then choose Δx based on its ability to resolve this sharp spatial change (Figure 3.2). While this figure shows that the differences between the step sizes are still subtle, it seems that the largest one, $\Delta x = 0.448$ mm, does not provide enough detail to show the sharp transition between cortex and electrode. The smaller step sizes appear to be more accurate and provide very similar solutions. Because it will be less computationally intensive to use

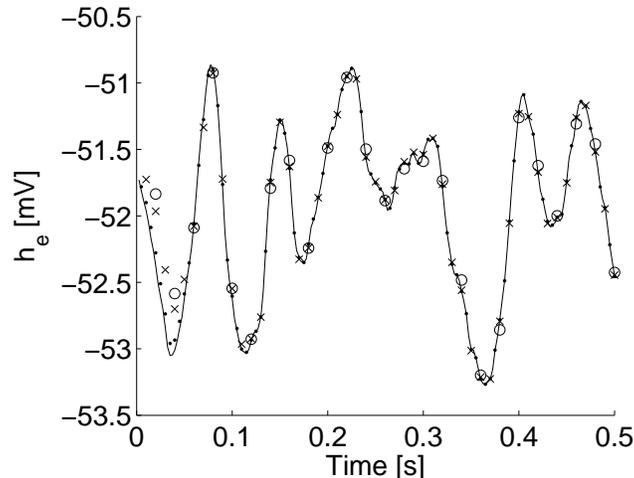


Figure 3.1: Convergence of numerical solutions as the time step is decreased using the method described in section 3.1.1. Here the spatial terms in the model have been removed to reduce it to an ODE. We used $\Delta t_0 = 2.5 \times 10^{-4}$ seconds, and we plotted the solutions for $\Delta t = 8\Delta t_0$, $4\Delta t_0$, $2\Delta t_0$, and Δt_0 . The data is represented by open circles, x markers, solid dots, and a solid line, respectively. The two smallest time steps (the solid line and solid dots) give overlapping results, indicating that the solution has converged. Note that, although the points at $8\Delta t_0$ do not accurately capture the peaks and troughs in the data, their absolute position is accurate. This study was done with typical excitation $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$.

$\Delta x = 0.224\text{mm}$ (or $\Delta \tilde{x} = 0.0008$ dimensionless), we choose this as the step size for our simulations.

In order to verify this in the typical case with no feedback control, we plot numerical solutions with $\Delta x = 0.448\text{mm}$, 0.224mm , and 0.112mm (Figure 3.3). Because the spatial solution appears to have converged at these step sizes, our choice of $\Delta x = 0.224\text{mm}$ is valid. The last task is to choose a value of Δt . We would like to use the largest time step for which the solution converges because this will result in the shortest computation time; this value is $\Delta \tilde{t} = 0.0001$ (dimensionless) or $\Delta t = 4 \times 10^{-6}$ seconds. Because we already showed that accurate solutions can be obtained with much larger values of Δt , this is an acceptable choice.

Therefore, the step sizes used in all of the following cortical simulations will be $\Delta x = 0.224\text{mm}$ and $\Delta t = 4 \times 10^{-6}$ seconds, based on adequate spatial resolution and stable time integration, respectively.

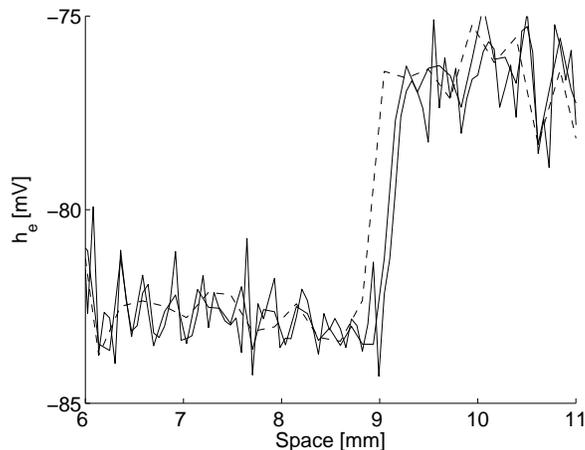


Figure 3.2: Convergence of numerical solutions over an electrode nonlinearity as the spatial step size is decreased using the method from section 3.1.1. The full simulation spanned 22.4mm, but here we show only 3mm of uncontrolled cortex and 2mm of cortex underneath an electrode while proportional feedback is applied. This figure shows solutions for $\Delta x = 4\Delta x_0$ (dashed), $2\Delta x_0$ (solid), and Δx_0 (solid), where $\Delta x_0 = 0.112\text{mm}$. Note that the two smallest time steps give very similar, overlapping solutions for \tilde{h}_e , while $4\Delta x_0$ gives an inaccurate result. Therefore, we choose the step size $\Delta x = 2\Delta x_0 = 0.224\text{mm}$ as a balance between accuracy and simulation cost. Other relevant parameters were $\Delta t = 2 \times 10^{-6}\text{s}$, $N = 80000$, $P_{ee} = 548.0$, and $\Gamma_e = 0.8 \times 10^{-3}$.

3.2 Data analysis techniques

3.2.1 Statistical analysis of EEG data

In chapter 5, our goal will be to identify a continuous range of sleep states by associating EEG data with a mathematical model of the human sleep cycle. To do this, it is necessary to characterize the EEG data using a set of features – this is similar to the way in which a sample of data points can be described by its mean and variance. For our analysis, we use a group of 17 different features. They are described below and implemented in Section 5.2.1.

Power in different frequency bands This group of five features consists of total power in the delta (up to 4 Hz), theta (4-7.5 Hz), alpha (7.5-12 Hz), beta (12-26 Hz), and gamma ranges (above 26 Hz).

Total power This is the total power in all five frequency bands.

Statistical measures These include variance, skewness, and kurtosis. Whereas the variance captures the spread of the data and is always positive, skewness is a measure of the asymmetry around the sample mean, i.e. negative skewness indicates that more data points lie below the mean than above. Kurtosis is a measure of how prone the distribution is to outliers; a signal with high kurtosis has infrequent large deviations from the mean.

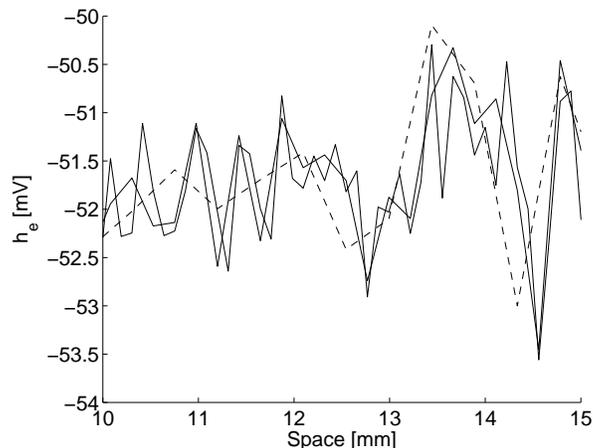


Figure 3.3: Convergence of numerical solutions as the spatial step size is decreased using the method described in section 3.1.1. Here we used $\Delta x_0 = 0.112\text{mm}$, and we plotted the solutions for $\Delta x = 4\Delta x_0$ (dashed), $2\Delta x_0$ (solid), and Δx_0 (solid). The two smallest time steps give very similar results, indicating that the solution has converged; this justifies our choice of step size, $\Delta x = 0.224\text{mm}$. This figure was created with typical excitation $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$, and $N = 80000$ time steps were used at $\Delta t = 2 \times 10^{-6}\text{s}$.

Spindle score The spindle detector identifies segments of the EEG signal where the difference between consecutive points changes from positive to negative five times in a row, thereby creating two peaks and two troughs. The lag parameter τ_L defines the number of sample points spanned by each rise or fall within the sought-for spindle, so it can be adjusted to search for these motifs at lower frequencies. We set $\tau_L = 5$, which allows for detection of 8-12Hz spindles in data sampled at 100Hz (with the maximum response occurring for spindles at 10Hz), and we used a minimum threshold of zero. The overall spindle score indicates the percentage of the signal that was classified as spindle activity. Matlab code for this function is provided in [31].

Permutation entropy Similar to the spindle score, the permutation entropy (PE) identifies motifs in the EEG data, such as peaks, troughs, and slopes. The PE has its maximum value when there is an equal distribution of all motifs and its minimum value when only a single motif is present. In this way, it is a measure of the “flatness” or “uncertainty” of the signal. Here we use the composite permutation entropy index (CPEI), which combines the PE with $\tau_L = 1$ and $\tau_L = 2$ with a minimum threshold level. The CPEI has been found to be a good measure of anesthetic depth [39]. In our study, we set the threshold at 1% of the inter-quartile range of the EEG data. Further descriptions of this measure and an associated MATLAB function can be found in [39].

Properties of log power These four features are based on the log of the power spectral density (PSD), as obtained by Welch’s method. First, we omit the delta and alpha peaks and calculate the slope and offset of a linear fit. We then determine the maximum value of the PSD above the linear estimate in the alpha range (8-17 Hz) and the

maximum value of the PSD in the delta range (0.5-4 Hz). These values will generally be large when a prominent peak is present. The code for generating these features was based on a Matlab function found in [24].

Power fractions The low power fraction is obtained by summing the power in the delta and theta ranges and dividing by total power. Similarly, the high power fraction is calculated by summation of the power in the beta and gamma ranges and dividing by total power.

3.2.2 Locally Linear Embedding (LLE)

Locally linear embedding is a method of nonlinear dimensionality reduction that was originally introduced in [45]. It is useful for visualizing high-dimensional data sets as they would be embedded in a low-dimensional space, and it can often uncover relationships and patterns that are masked by the complexity of the original data set. It has been used to obtain maps of facial expressions and classify handwritten digits [48], as well as discriminate between normal and pre-seizure EEG measurements [2]. In this dissertation, we use LLE to characterize sleep EEG data and the numerical solutions of the cortical model. By embedding both in a two-dimensional space, we will be able to associate traditional EEG sleep stages with the continuous spectrum of states provided by the model.

The algorithm

Let us begin with a high-dimensional data set stored in a matrix \vec{X} of size $D \times N$, where each column \vec{X}_i represents one of the N D-dimensional data points. Then the LLE algorithm consists of three steps:

1. **Calculate the nearest neighbors of each data point \vec{X}_i in the D-dimensional space.** This can be done in several ways; for example, we might choose the k closest points based on Euclidian distance, or we may choose only the points within a sphere of a given radius.
2. **Determine the best reconstruction of each point using only its nearest neighbors.** Mathematically, this takes the form of a least squares minimization problem:

$$\min_W \sum_{i=1}^N \left| \vec{X}_i - \sum_{j=1}^k W_{ij} \vec{X}_j \right|^2, \quad (3.3)$$

where k represents the number of nearest neighbors. Our goal is to choose the weights W that best reconstruct the original data points in the D-dimensional space, based on the criteria of least-squared error. Because we use only the nearest neighbors, we must have $W_{ij} = 0$ if \vec{X}_j is not a neighbor of \vec{X}_i . In addition, we guarantee invariance to translations by enforcing $\sum_j W_{ij} = 1$. Note that the minimization can be calculated individually for every i .

3. **Compute the low-dimensional output vectors \vec{Y}_i .** These are chosen to provide the best global reconstruction using the weights W from the previous step. Again, this can be formulated as a least squares minimization:

$$\min_{\vec{Y}} \sum_{i=1}^N \left| \vec{Y}_i - \sum_{j=1}^k W_{ij} \vec{Y}_j \right|^2 . \quad (3.4)$$

Here we are making the assumption that the weights that give the best reconstruction in D dimensions will also be the optimal weights in the lower-dimensional space. In this case, the N minimization problems are coupled by the elements of \vec{Y} , so they must be solved simultaneously.

A detailed description of the algorithm and several examples are provided in [48]. In addition, a Matlab implementation of LLE is available on the authors’ website [46]; it was used to generate all results presented here.

As a simple example, consider using LLE on a known 3D manifold (Figure 3.4). In this toy example, the underlying manifold is known (although normally this would not be the case), and we recognize that it has only two dimensions, despite living in 3-dimensional space as shown in Subfigure 3.4(a). The data set \vec{X} consists of a random sampling of points from the manifold, as in Subfigure 3.4(b), and the LLE output for a reduction to two dimensions is displayed in Subfigure 3.4(c). Here we see that LLE successfully unravels the manifold and uncovers its true 2D nature.

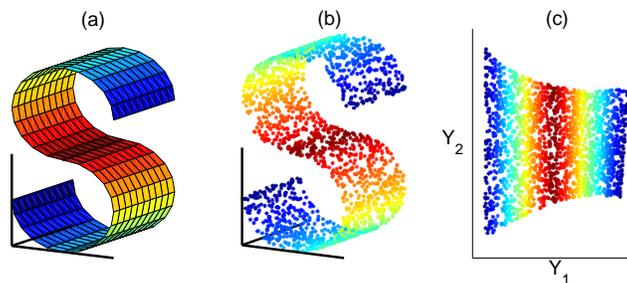


Figure 3.4: A simple example of LLE, where three dimensions are reduced to two. (a) The underlying manifold, which lives in 3D space but has only two dimensions. In a typical LLE problem, the shape of this manifold is unknown and has too many dimensions to visualize easily. (b) A sampling of points from the manifold, which serves as the input to the LLE algorithm. (c) The result of applying LLE to the data set in (b). Note that in the $Y_1 - Y_2$ output space, the manifold has been flattened to reveal its two principal dimensions. This figure was generated using the “scurve.m” code from the LLE website [46].

A possible source of confusion with locally linear embedding is the interpretation of output dimensions such as Y_1 and Y_2 . Unlike linear methods such as principal component analysis, LLE does not provide a description of the output vectors in terms of the original D dimensions. The elements of \vec{Y} are chosen to give the best *local* reconstructions based on a global minimization problem; this means that the interpretation of \vec{Y} is different for every data point, and it cannot be described by a simple combination of the original dimensions.

3.2.3 Identification of strongly connected components

The use of the LLE algorithm is based on the assumption that the entire data set lies on the same manifold in high-dimensional space. If more than one manifold is present, the locally linear reconstructions will no longer be accurate (imagine, for example, a point with nearest neighbors located on two separate manifolds). Therefore, before using LLE on a data set, we must verify this assumption.

The mathematics and terminology of directed graphs allows us to accomplish this task [59]. Note that when we calculate the nearest neighbors in the first step of the LLE algorithm, we create a directed graph based on the data points. For example, suppose there is a data set of seven points, and we have determined that point 2 is a neighbor of point 1, point 5 is a neighbor of point 2, etc. This can be depicted by arrows drawn from each point to its neighbors (Figure 3.5). Then we can define a *strongly connected component* as a group of points where the arrows created by nearest neighbor associations allow for travel from every point in the group to every other point in the group [59]. *When a group of data points is strongly connected, this indicates that they lie on the same manifold* [48].

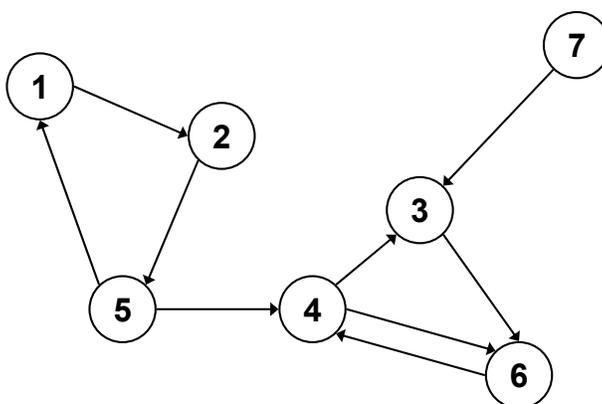


Figure 3.5: An example of a directed graph generated by nearest neighbor associations. Here point 2 is a neighbor of point 1, point 3 is a neighbor of point 7, etc. In this case, the directed graph forms two strongly connected components: points 1, 2, 5, and points 3, 4, 6. In analyzing this data set, we would use LLE separately on each of these components and would remove point 7, which is not strongly connected to any other point.

The example in Figure 3.5 has two strongly connected components: points 1, 2, 5, and points 3, 4, 6. However, the two groups are not strongly connected together; one can move from the first group to the second through the connection between 4 and 5, but there is no way to get from the second group to the first. Point 7 is not strongly connected to any other point. Therefore, to use LLE on this sample data set, we would remove point 7 and use the algorithm separately on each strongly connected component.

There are several ways to identify the strongly connected components of a data set. The most traditional method involves an algorithm based on depth-first search of the directed graph [59]. An alternative method relies on analysis of the eigenspace that results from the LLE calculations [41]. It is also true that choosing the nearest neighbors in a different manner

or increasing the value of k can change the structure of the strongly connected components. However, for the purposes of this study, we used the MATLAB function `dmperm` on a matrix containing the nearest neighbor associations for the data set. This function, based on the Dulmage-Mendelsohn decomposition, permutes the rows and columns of a matrix to put it into block diagonal form; by including the fact that every point is a neighbor with itself, we can guarantee that this permutation will be symmetric. As output, `dmperm` provides the new order of rows and columns and identifies the blocks of the permuted matrix, where each block represents one strongly connected component within the data.

Chapter 4

Feedback control for the suppression of epileptic seizures

We previously used the mathematical model of the cortex presented in Chapter 2 to study feedback control of epileptic seizures. We first showed that linear (proportional) feedback, differential control, and feedback based on a filtered version of h_e could suppress the seizure-like waves in the model [22]. This was done by measuring and controlling at every single point in space, with the assumption that the measurement of a cortical surface electrode was proportional to the mean soma potential, h_e . We later added the capability to define the number, size, and spacing of the electrodes used in the simulation [29]. We assumed that the potential at every point was proportional to h_e and that the single measurement of an electrode was equal to the mean of the potential for all points lying below it.

In this chapter, we incorporate the new step sizes discussed in chapter 3 and present three additional improvements: the ability to incorporate a detailed profile for electrode effectiveness, a new model for the measurement of a cortical surface electrode, and a method for performing charge-balanced feedback control. We use the formulation of the model from Section 2.2.1 with the parameters listed in Table 2.1.

4.1 Refinements to the model of feedback control

4.1.1 Electrode profile

The use of smaller increments in space (resulting from our study of numerical convergence in Section 3.1.1) allows us to make a further improvement to the simulations of feedback control: the addition of a more detailed electrode profile. We previously assumed that the electrodes maintained a constant profile across their surface while measuring or applying the stimulus (i.e. every point on the electrode sensed or provided the same value) and that no tissue beyond the edge of the electrode was affected by its activity. However it has been shown experimentally that this is not the case [56]. Here, we take the first step towards a realistic electrode model by including a smooth falloff at the electrode edges. The falloff is incorporated into the function $a(x, t)$, so that the gain varies between zero (over uncontrolled cortex) and a_{\max} (under the electrodes) via a hyperbolic tangent function. This gain function

is used in the application of control to indicate that the influence of the electrode decreases with distance, and it is also applied during sensing with $a_{\max} = 1$ to indicate that the cortical tissue has less impact on the electrode measurement as distance increases. We have not yet included any variation over the surface of the electrode, but this is certainly an adjustment that can be considered in the future.

4.1.2 Model of electrode measurements

Our previous simulations of feedback control [22, 29] utilized two key assumptions: 1) the signal measured by an electrode on the cortical surface is proportional to h_e , the mean soma potential of the excitatory neuron population, and 2) a voltage applied to the surface of the cortex via electrode directly affects the average soma voltage in that region. The first assumption allows us to define the control effort u in terms of \tilde{h}_e (in this case, \tilde{h}_e would represent the measured voltage), and the second assumption implies that the expression for u can be added directly to the SPDE model in equation (2.1). While the latter assumption appears to be valid, there is evidence that we cannot write u as an explicit function of \tilde{h}_e as the first assumption suggests. It is likely that the voltage sensed by a surface electrode is different than the averaged soma voltage, \tilde{h}_e .

First, it is important to realize that the signal measured by an electrode is a function of the extracellular *currents* in the tissue, rather than the intracellular somatic potential [38]. We define the signal sensed at a point on the cortical surface to be \tilde{h}_m . Then, to understand the difference between \tilde{h}_e and \tilde{h}_m , we consider a pyramidal neuron in the cortex with one excitatory synapse as shown in Figure 4.1. Suppose that the pyramidal neuron receives excitatory input due to a proximal synapse in layer 4 (Figure 4.1a); this will cause intracellular flow of ions that will induce a current dipole with sources (+) on the apical dendrite near the surface and sinks (-) near the soma. The surface electrode \tilde{h}_m will sense the extracellular current source near the surface and will thus depolarize. The soma potential \tilde{h}_e will also depolarize due to the excitatory input; therefore, in this case, both \tilde{h}_e and the surface electrode show a depolarization. On the other hand, suppose that the pyramidal neuron receives excitatory input due to a distal synapse in layer 1 (Figure 4.1b). Because the input is still excitatory, the neuron will depolarize, and this will be reflected in the soma potential \tilde{h}_e . However, this input will cause an extracellular current dipole with reverse polarity; there will be a source (+) near the soma and current sinks (-) near the surface. This means that the voltage sensed by the surface electrode will show a hyperpolarization. Therefore, in this case, the deflection of \tilde{h}_e and the signal seen by the surface electrode are different [18].

This implies that we should no longer use \tilde{h}_e as the measured electrical potential in our expression for the control effort u . Instead, the measurement will be a function of the currents in the cortex due to synaptic inputs, denoted as \tilde{h}_m . We will refer to this as the sensed signal. Then, for the purposes of feedback control, the applied effort u will be a function of \tilde{h}_m .

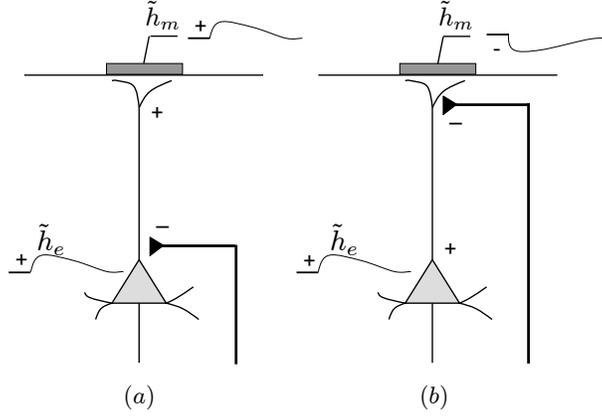


Figure 4.1: Relationship of the sensed signal \tilde{h}_m to the mean soma potential \tilde{h}_e . In (a), an excitatory input synapses near the soma, causing a depolarization in \tilde{h}_e . Similarly, the dipole current generated by this synaptic event involves current sources near the cortical surface, which are manifested as a depolarization in the electrode measurement \tilde{h}_m . However, when the excitatory input occurs near the surface as in (b), the current dipole is reversed, causing opposite deflections in \tilde{h}_e and \tilde{h}_m . This figure is modeled after Box 46-1 in [18].

Basic form of model

To determine the composition of the sensed signal, we must consider the extracellular current flows due to three types of synaptic input (local intracortical input, long-range corticocortical input, and subcortical input). To do this, we need to know whether the inputs are excitatory or inhibitory and whether they synapse near the surface or near the soma. In what follows, we take care to distinguish depolarization in the electrode measurement from depolarization of the soma.

- **Local intracortical inputs, $N_e^\beta \tilde{S}_e$ and $N_i^\beta \tilde{S}_i$.** Within a cortical macrocolumn, excitatory synapses tend to occur close to the surface, while inhibitory synapses are located near the soma [38, 51]. Thus, the excitatory inputs are as depicted in Figure 4.1b, and the inhibitory inputs have the geometry of Figure 4.1a but with the opposite sign (because Figure 4.1 depicts *excitatory* inputs). Each of these configurations will cause a hyperpolarization in the electrode measurement; therefore, both terms will have negative signs in the measurement model: $-AN_e^\beta \tilde{S}_e - BN_i^\beta \tilde{S}_i$, where A and B are positive constant weights to be determined.
- **Long-range corticocortical input, $\tilde{\phi}_e$.** Corticocortical inputs are exclusively excitatory [55, 26, 38] and tend to synapse near the surface [38, 51]. More specifically, layers 2 and 3 of the cortex seem to have a higher density of corticocortical inputs [18, 37]. As shown in Figure 4.1b, this input type will cause a hyperpolarization in the electrode signal and will thus have a negative sign in the measurement model: $-C\tilde{\phi}_e$, where C is a positive constant weighting factor to be determined.

- **Subcortical inputs, $(P_{ee} + \tilde{\Gamma}_1)$ and $(P_{ie} + \tilde{\Gamma}_3)$.** While the distribution of these synapses is not clear-cut, it seems that subcortical inputs terminate most densely in layer 4 near the soma [18, 37]. Because P_{ee} is an excitatory input of the type shown in Figure 4.1a, it will have a depolarizing effect on the electrode measurement; therefore, we give it a positive sign: $+D(P_{ee} + \tilde{\Gamma}_1)$, where D is a positive constant weight. On the other hand, P_{ie} is inhibitory and will thus have a hyperpolarizing effect: $-E(P_{ie} + \tilde{\Gamma}_3)$, where E is a constant weighting factor. The values of D and E are to be determined.

Incorporating all three types of synaptic input gives us this basic expression:

$$\text{measured current} \sim -AN_e^\beta \tilde{S}_e - BN_i^\beta \tilde{S}_i - C\tilde{\phi}_e + D(P_{ee} + \tilde{\Gamma}_1) - E(P_{ie} + \tilde{\Gamma}_3), \quad (4.1)$$

with A, B, C, D, E of positive sign but (so far) unknown magnitude. The consequence of these inputs is only evident after synaptic transmission. Therefore, we include a rate constant for this process by using an equation similar to that of \tilde{I}_{ee} in the SPDE model. Let \tilde{I}_m represent the current measured at the cortical surface and T_m represent a rate constant. Then

$$\left(\frac{1}{T_m} \frac{\partial}{\partial t} + 1 \right)^2 \tilde{I}_m = F \left(-AN_e^\beta \tilde{S}_e - BN_i^\beta \tilde{S}_i - C\tilde{\phi}_e + D(P_{ee} + \tilde{\Gamma}_1) - E(P_{ie} + \tilde{\Gamma}_3) \right), \quad (4.2)$$

where A, B, C, D, E , and F are positive constant weights. We choose $T_m = 12.0$ to match the rate constant of the excitatory population T_e . The values of A through E will depend on the number of synapses of each type and the average distance of the synapse from the soma. The coefficient F is a gain parameter that will scale the magnitude of all the synaptic inputs; this ensures that they have the appropriate amount of influence over the electrode measurement \tilde{I}_m . In addition, we can think of F as containing the effective resistance of the cortex. Recall that the electrode measurement is determined by *currents* in the cortex, yet the components on the right side of (4.2) are based on voltages. Because the currents produced by these voltages can be calculated with Ohm's Law [18], the gain parameter F provides the necessary conversion.

To complete the model of the electrode measurement \tilde{h}_m , we must account for the reversal potential of the cortical neurons. This determines the direction of current flow associated with the inputs described above (we previously assumed that the neurons were at resting potential). We once again take our cue from the SPDE model and define

$$\tilde{h}_m \equiv (h_0^e - \tilde{h}_e) \tilde{I}_m = \left(\frac{45 - h_e}{-70} \right) \tilde{I}_m. \quad (4.3)$$

Thus, (4.2) and (4.3) comprise a complete model of the potential sensed by a cortical surface electrode, \tilde{h}_m . In our simulations of feedback control, the applied electric field u will be a function of this variable.

Estimation of coefficients

We have not yet assigned numerical values to the coefficients A through E . To do this, we think of them as the percentage of pyramidal neuron synapses due to each source. For

Coefficient	Input type	Lopour/Szeri value	Liley/Wright value
A	local excitatory	0.413	0.324
B	local inhibitory	0.092	0.088
C	long-range excitatory	0.458	0.583
D	thalamocortical excitatory	0.034	0.006
E	subcortical inhibitory	0.004	0.000

Table 4.1: Values of the coefficients for equation (4.2). The first column of values was estimated as described here, while the second column was derived using probabilistic methods in [28].

example, A will represent the percentage of synapses on any given pyramidal neuron that come from other excitatory neurons in the same macrocolumn. There are three physiological relationships that allow us to determine these values:

1. The number of synapses on pyramidal cells due to local cortical neurons is roughly equal to the number of synapses due to cortical neurons in other macrocolumns or in the contralateral hemisphere [1]. This implies that $A + B = C$.
2. Approximately 98 percent of synapses on pyramidal cells are corticocortical, while 2 percent are thalamocortical [1, 38]. This implies that $A + B + C = .98$ and $D + E = .02$.
3. Roughly 90 percent of all cortical synapses are excitatory and 10 percent are inhibitory [1, 3]. This implies that $A = 9B$ and $D = 9E$.

After solving the above equations, we account for the fact that synapses near the soma will have a greater influence on the electrode measurement [38] by multiplying B , D , and E by a factor of two.

While this method of estimation may seem unsophisticated, others have achieved similar results through more detailed probabilistic analysis [28]. Both are listed in Table 4.1 for comparison; note that each set of coefficients has been scaled to add to 1. In simulation, when comparing the two sets of coefficients A through E , the only difference in the resulting \tilde{h}_m appears to be an offset. Because offsets are not reflected in EEG measurements, this difference is inconsequential. Hence, two completely independent approaches provide essentially equivalent coefficients for the sensed signal.

Verification of full model

We can verify our model by returning to the hypothetical pyramidal neuron described at the beginning of Section 4.1.2 and based on Box 46-1 in [18]. As before, suppose that we are modeling the electrode measurement of a pyramidal neuron with excitatory inputs in both cortical layers 1 and 4. If a majority of the inputs occur in layer 1 near the cortical surface as in Figure 4.1(b), \tilde{h}_e and \tilde{h}_m will have similar dynamics, but a hyperpolarization in one signal will be a depolarization in the other; the signals will be negatively correlated. This behavior is seen in our model for \tilde{h}_m . If we run a simulation with the typical set of parameters (where local and corticocortical connections dominate because P_{ee} is low), we

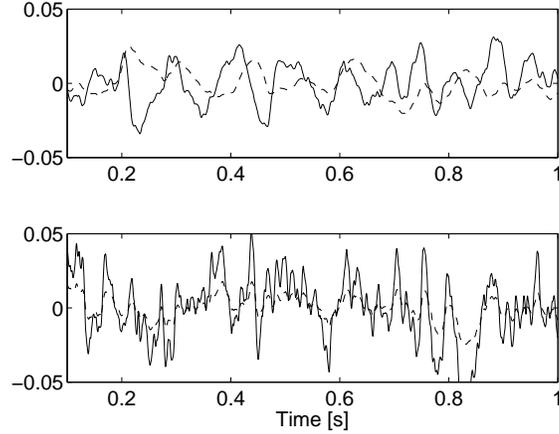


Figure 4.2: Comparison of \tilde{h}_m (solid) and \tilde{h}_e (dashed) with two sets of parameters. At typical levels of excitation where $P_{ee} = 11.0$ and $\Gamma_e = 1.42 \times 10^{-3}$, the two signals are negatively correlated (top subfigure), as predicted by our physiological analysis. This negative correlation is especially noticeable if the synaptic time delay is taken into account by shifting \tilde{h}_e slightly to the left. However, when we simulate a strong excitatory input near the soma by setting $P_{ee} = 1000.0$, the signals become positively correlated (bottom subfigure). This was also predicted by the physiology of the cortex, and thus helps justify our model of \tilde{h}_m . Note that for both sets of parameters, the signal offsets (means) were removed to facilitate a direct comparison.

see that \tilde{h}_e and \tilde{h}_m are negatively correlated (top of Figure 4.2). On the other hand, when the strongest input is near the soma in layer 4 as in Figure 4.1(a), it will have the same effect on both signals, and they will be positively correlated. This, too, is demonstrated by the measurement model. We can simulate a large excitatory input near the soma by increasing the value of P_{ee} ; when we run the simulation with this change, we see that \tilde{h}_e and \tilde{h}_m become positively correlated (bottom of Figure 4.2). Thus, the measurement model accurately reproduces the physiological effects of varying cortical inputs, and we will use it in subsequent simulations of feedback control. The applied electric field u will be a function of \tilde{h}_m as opposed to \tilde{h}_e .

For reference, we compare \tilde{h}_e and \tilde{h}_m at seizure parameters ($\Gamma_e = 0.8 \times 10^{-3}$ and $P_{ee} = 548.0$) in Figure 4.3. In this case, the signals have a large negative correlation, but very similar dynamics. This suggests that it will be possible to suppress seizures with feedback control based on \tilde{h}_m (as it was with \tilde{h}_e in [22]), although we may need to use a gain of the opposite sign.

4.1.3 Charge-balanced control

In choosing a function to represent the applied electric field u , we start with the concept of proportional feedback control. This is the simplest and most common type of control – intuitively, the applied effort should be proportional to the error between the measured

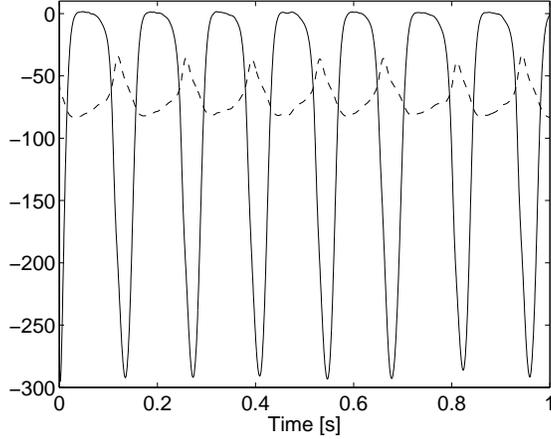


Figure 4.3: Comparison of the electrode measurement h_m (mV, solid) and the mean soma potential h_e (mV, dashed) at levels of subcortical excitation that cause seizure-like oscillations, $P_{ee} = 548.0$ and $\Gamma_e = 0.8 \times 10^{-3}$. Here we see that the two signals are negatively correlated, but have very similar dynamics. This indicates that it should be possible to perform feedback control using our new measurement model \tilde{h}_m (as we did previously with \tilde{h}_e in [22]), although a positive gain may be necessary.

signal and its desired value. In this case, we can define proportional control as

$$u = a(x, t)(\bar{h}_m + b) , \quad (4.4)$$

where $a(x, t)$ is the control gain and \bar{h}_m is the measured electrode potential; it is calculated by taking the spatial average of \tilde{h}_m (4.3) under each electrode. The parameter b is a constant offset that can be tuned to achieve the desired equilibrium value of \tilde{h}_e . While this control algorithm is able to suppress the seizure-like oscillations of the model, it would be difficult to implement safely. Whenever stimulation is applied, it is important that the process be both charge-balanced and chemically reversible in order to prevent damage due to the production of new chemical species. There is a threshold for reversibility called the “reversible charge injection limit,” which signifies the maximum allowable charge injection before the polarity is reversed [43]. Because a proportional controller does not penalize the amount of effort used (i.e. the magnitude of the applied electric field), it relies on large signals of only one sign, which would exceed this threshold over time. The chemical processes associated with this type of stimulation would therefore be irreversible and damaging to cortical tissue.

To improve on this method, we may consider adding a derivative or integral component to the controller, or even using all three terms to create a proportional-integral-differential (PID) controller [9]. The derivative term increases or decreases the control effort based on the rate of change of the error. This can reduce the response time of the controller because the derivative term “anticipates” the behavior of the system. For example, a control law incorporating both proportional and differential control might take the form

$$u = a(x, t) (\bar{h}_m(t) - \bar{h}_m(t - \tau_d)) , \quad (4.5)$$

so the applied signal is proportional to the difference between the electrode measurement now and the electrode measurement at a previous time [22]. The time delay is set by the parameter τ_d .

Because this differential controller utilizes the same harmful voltages mentioned previously, we choose to implement a controller with an integral term:

$$u = a(x, t)(\bar{h}_m + b) + c(x, t) \int u dt . \quad (4.6)$$

Here, $c(x, t)$ is another gain term. It will be negative, meaning that this new term will oppose the total integral of the applied voltage u . If the integral of u is positive, it will add a negative component to the applied voltage, and vice versa; in this way, it pushes the integral of u to zero. In other words, it forces the applied signal to be charge-balanced and thus safe for cortical tissue. Because we have also included the proportional control term, this feedback setup will still suppress the seizure oscillations. Note that this is different than traditional integral control, which is based on the integral of the *error* between the desired value of the signal and its actual value.

4.2 Simulations of feedback control

We are now in a position to simulate the suppression of seizure waves using various control methods: proportional control with or without offset, differential control, and total integral feedback control. We include all the improvements discussed in the previous section. The results shown here were generated using an Intel Core 2 Duo 2.13 GHz processor, and the calculations took roughly 5-15 minutes at the highest spatial and temporal resolutions (depending on the method of control, the magnitude of the gain, and the number of electrodes). The code was written and executed in MATLAB.

4.2.1 Proportional control

Figure 4.4 shows the result of applying proportional control from equation (4.4) to a simulated seizing cortex. In the first half of the time interval, we see a seizure wave emanating from the center of the space, just as in Figure 2.3; then, at 0.25s the control is switched on, and the pathological seizure activity quickly disappears. In this case, the feedback stimulation is applied via five electrodes that are 11.2mm across, with gain $a_{\max} = 2$ and $b = 0$. The geometry of these electrodes can be clearly seen in Figure 4.4(b), which shows the value of the applied signal u [mV] for the same simulation. Here, we see that the control effort is zero for the first half of the time interval, but when the control is turned on, each electrode applies a negative potential. The signal with the largest magnitude comes from the electrode at the center of the simulation space, where P_{ee} is highest. Note that it is not necessary to have electrodes covering the entire length of the seizure wave. When we place the electrodes at the center of the “hot spot,” they are able to halt the outward motion of the wave. For more information on the parameters used to create the figures, please refer to the figure captions.

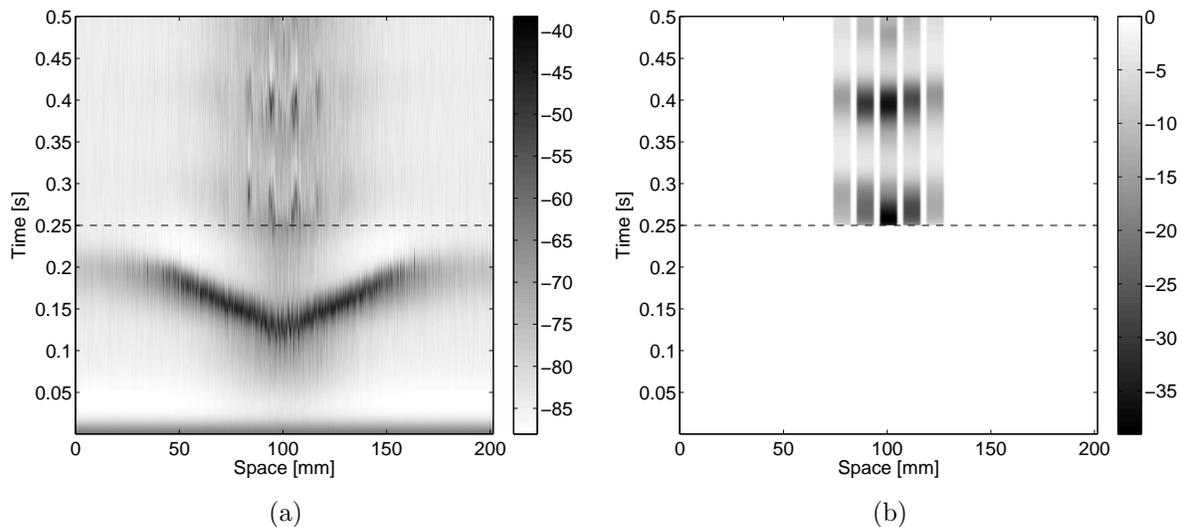


Figure 4.4: Results for proportional control. (a) Plot of h_e [mV] in space and time. In the first 0.25 seconds, we see a seizure wave traveling on the simulated cortex because we have set the excitation parameters to $P_{ee} = 548.0$ and $\Gamma_e = 0.8 \times 10^{-3}$. At 0.25 seconds (dashed line), the proportional controller (4.4) is turned on, and we see that the seizure-like waves are immediately suppressed. For this simulation, we used \tilde{h}_m as the electrode measurement, and control was applied via five independent electrodes of width 11.2mm with a profile defined by the hyperbolic tangent function. The controller gains were $a_{\max} = 2$ and $b = 0$. (b) Plot of the control effort u [mV] associated with the simulation in (a). This shows the geometry of the five electrodes; in this case, they apply only negative signals. Note that, for clarity, the grayscale in (b) is the opposite of the one used in (a).

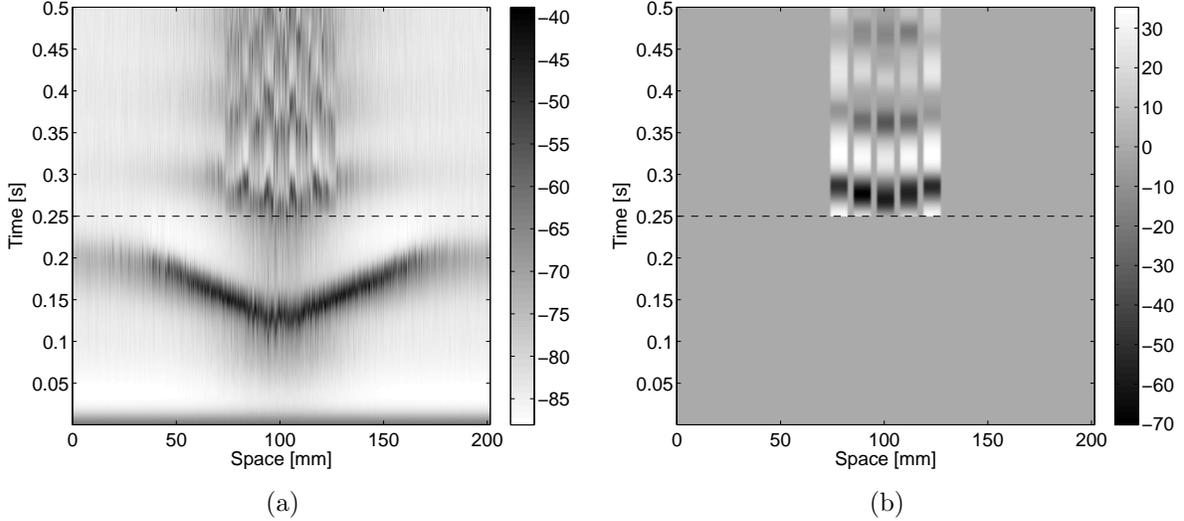


Figure 4.5: Results for a proportional controller with non-zero offset. (a) Plot of h_e [mV] in space and time. All parameters and settings are the same as in Figure 4.4, except $b = -0.3$. Again, the seizure-like waves are suppressed. In this case, the remaining fluctuations of h_e are larger than they were for proportional control (Figure 4.4), but now the controller is applying both positive and negative potentials. (b) Plot of applied voltage u [mV] for the five electrodes used in the simulation.

We can alter these results by using a non-zero offset b . Figure 4.5 shows a simulation of proportional feedback control with $a_{max} = 2$ and $b = -0.3$. This change in offset causes the controller to push toward a different equilibrium point, as is evidenced by the darker color underneath each of the electrodes as compared to Figure 4.4. With the control turned on, the fluctuations of h_e are larger than they were with $b = 0$, but this can be optimized by jointly tuning the values of a_{max} and b . Here we have used the same a_{max} to facilitate comparison between $b = 0$ and $b \neq 0$. The most important change is that, for $b = -0.3$, the electrodes apply both positive and negative signals. This is shown in Figure 4.5(b). While these signals are not truly charge-balanced, they do present an improvement over those resulting from traditional proportional control.

4.2.2 Differential control

Another strategy for feedback control is differential control, where the applied signal is proportional to the difference between a measurement at the current time and a measurement at a previous time. This is stated mathematically in equation (4.5). Results of applying this control law are shown in Figure 4.6. Again, the application of feedback control is successful in counteracting the seizure-like oscillations. As with offset proportional control, the applied signals oscillate between positive and negative values; however, they are not explicitly charge-balanced. Another interesting feature of this control strategy is that each electrode applies a signal with a frequency greater than that of the seizure wave. This is most likely due to our choice of τ_d and the uncontrolled frequency of the seizure oscillation.

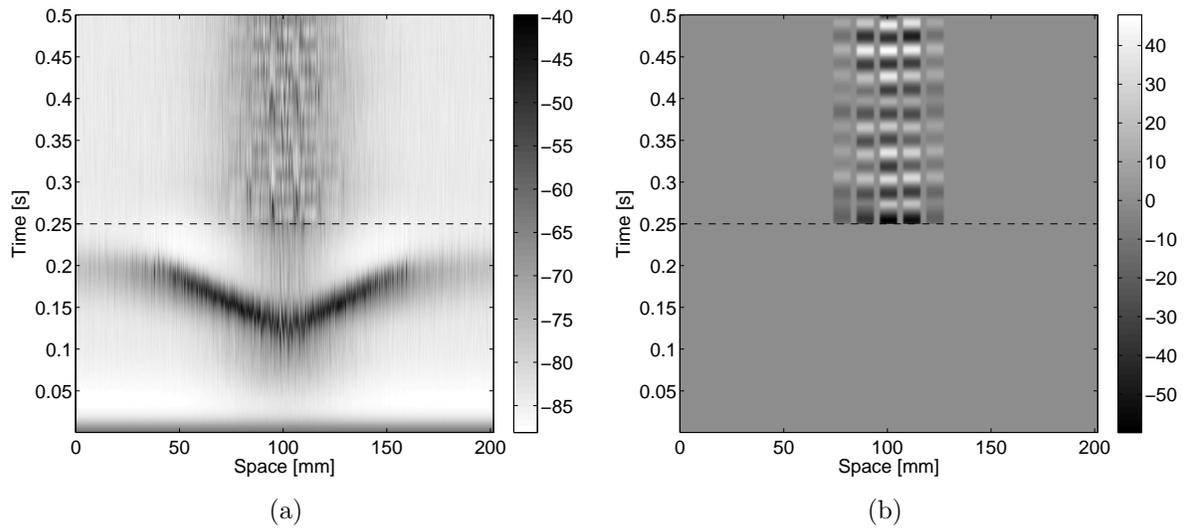


Figure 4.6: Results for differential control. (a) Plot of h_e [mV] in space and time. Excitation parameters and electrode geometry are the same as those used in Figure 4.4. At 0.25 seconds (dashed line), the differential controller (4.5) is turned on, and the seizure-like oscillations are suppressed. Again, the controller applies both positive and negative potentials; however, in this case, the frequency is higher than that of the seizure waves. The controller gains were $a_{\max} = 5$, $b = 0$, and $\tau_d = 20$ ms. (b) Plot of applied effort u [mV] for the simulation of differential control.

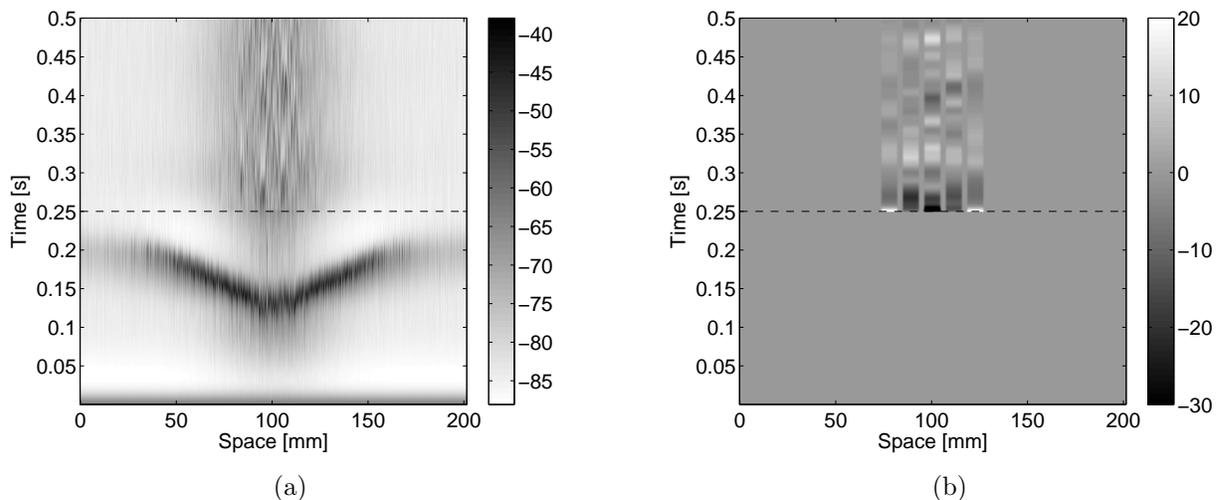


Figure 4.7: Results for total integral feedback control. (a) Plot of h_e [mV] in space and time. Simulation parameters and electrode geometry are the same as those in Figure 4.4. The controller gains were $a_{\max} = 8$, $b = -0.1$, and $c = -8$. (b) Plot of applied effort u [mV] in space and time, corresponding to the simulation of feedback control in (a). The potential applied by the electrodes is zero until $t = 0.25$ seconds, when the controller is turned on (dashed line). Then each electrode oscillates between positive and negative signals, indicative of the “charge-balanced” nature of the control. For this simulation, the average magnitude of net applied effort over all five electrodes was 0.23mV. This value will approach zero as t increases, and it can also be reduced by increasing the integral control gain, c .

4.2.3 Total integral feedback control

Figures 4.7 and 4.8 show the effect of total integral feedback control from equation (4.6) on the seizure-like behavior of the model. Again, each electrode varies between positive and negative voltages once the controller is turned on. However, in this case, the total signal applied by each electrode is very close to zero. This is the desired result because it indicates that each electrode applies a balanced signal. In addition, Figure 4.8 indicates that, with the exception of the initial spike in control effort to -60 mV at $t = 0.25$ s, the applied potentials are very small. The overall magnitude is roughly 20mV, compared to 90 – 100mV for differential control or offset proportional control. With appropriate timing for the start of the control relative to the seizure wave, it may be possible to reduce this initial spike in effort.

Thus, in Figures 4.4 through 4.8 we have demonstrated that the new model for electrode measurements h_m can be used to suppress seizure waves via feedback control. Further, if the controller u contains the integral of the total applied effort, then control can be applied in a manner that is thought to be safe for cortical tissue.

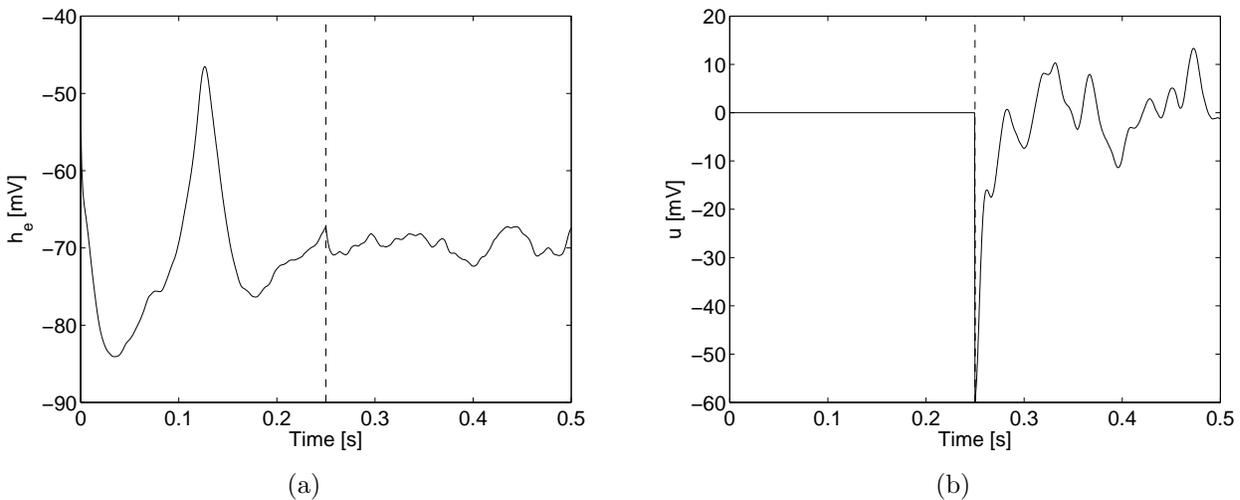


Figure 4.8: Results for total integral feedback control at a single point in space. (a) Plot of h_e in time extracted from Figure 4.7(a) at 100.8mm. (b) Plot of control effort u in time extracted from Figure 4.7(b) at 100.8mm. Here we see that the control effort oscillates between positive and negative values in a charge-balanced manner.

4.3 Spatial properties of the feedback control model

The specifications for successful seizure suppression via feedback control are inherently dependent on the physiological characteristics of the cortex. We have already mentioned the need for charge-balanced signals to avoid damaging the tissue, and extracellular electrochemical changes must be taken into account during electrode design. Another cortical property that has implications for feedback control is the length scale for falloff of connections between neurons. This parameter tells us how electrical signals are dispersed through the cortex. Here we discuss how it relates to the spread of seizure waves and to the electrode size necessary to stop their propagation. This can be regarded as an illustration of the utility of having a mathematical model; basic design tradeoffs can be explored.

4.3.1 The cortical length scale

In the current mesoscale cortical model, the parameter λ_e represents the *inverse* length scale for corticocortical connections. Therefore, an increase in λ_e corresponds to shorter, denser connections between neurons. A decrease in λ_e means that a neuron will have synapses over a greater distance, but its connections will be weaker. For all simulations up to this point, we have used $\lambda_e = 11.2$ (dimensionless), which corresponds to a length scale of 25mm. The role of λ_e can be understood by examining its effect on the oscillations in the model. Figure 4.9 shows seizure waves for three different spatial length scales. Subfigure 4.9(b) uses the default value. When we divide this by two (thereby increasing the length scale), the seizure waves spread further along the model cortex, but have a smaller amplitude; conversely, when we double the typical value of λ_e , the waves have a slightly higher amplitude and cover a shorter length of cortex. These two cases are depicted in Subfigures 4.9(a) and

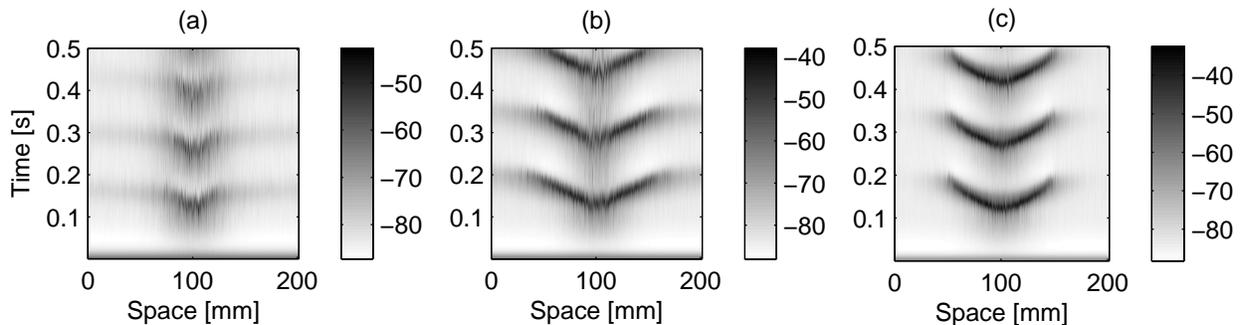


Figure 4.9: Seizure waves for three values of the inverse corticocortical length scale λ_e : (a) 5.6, (b) 11.2, and (c) 22.4. An increase in λ_e corresponds to shorter, denser connections between neurons; this results in slower, higher amplitude waves over a shorter length of cortex. As with previous simulations, we used $\Gamma_e = 0.8 \times 10^{-4}$, $\alpha = 1.6$, and a gaussian distribution of P_{ee} with a maximum value of 548.0 at the center of the interval.

(c), respectively. We also notice that the wave speed decreases with λ_e .

4.3.2 Implications for electrode design

Based on Figure 4.9, it is reasonable to assume that the electrode size needed to control seizure waves may be related to the value of λ_e . For example, we might expect that dense, compact oscillations with a low wave speed might require smaller electrodes with closer spacing than waves that are diffuse and travel quickly.

We can examine this hypothesis using the mesoscale cortical model and our formulation of feedback control. For spatial uniformity and ease of interpretation, we will use a constant P_{ee} across the whole interval. This creates continuous traveling waves, such as those depicted in Subfigure 4.10(a). This simulation used $\lambda_e = 11.2$ and produced waves with a frequency of approximately 9Hz and a wave speed of roughly 3m/s. These properties vary slightly with λ_e ; for example, $\lambda_e = 5.6$ leads to 11Hz waves that travel at slightly more than 3m/s, and $\lambda_e = 22.4$ creates 8.5Hz waves with a speed of approximately 2.6m/s. For each value of the spatial length scale, we will add a single electrode to the center of the interval and perform proportional feedback control; an example of this is shown in Subfigure 4.10(b). By comparing simulations with different values of λ_e and varying electrode sizes, we can characterize their relationship to control effectiveness.

First, however, we must define what we mean by “effective” control. In some simulations of feedback control, it is clear that the applied signal is successfully counteracting the seizure waves (for example, Figures 4.4 and 4.7). In other cases, such as those in Figures 4.5 and 4.6, the controller seems to perform in a non-optimal way. How can we characterize the difference between these results? Measures such as the amplitude histogram and the Teager energy have previously been used for comparing controlled and uncontrolled seizures in rats [5]. In this case, we use the standard deviation as a simple way of quantifying the effect of the control. We have seen that uncontrolled seizure waves have a standard deviation on the

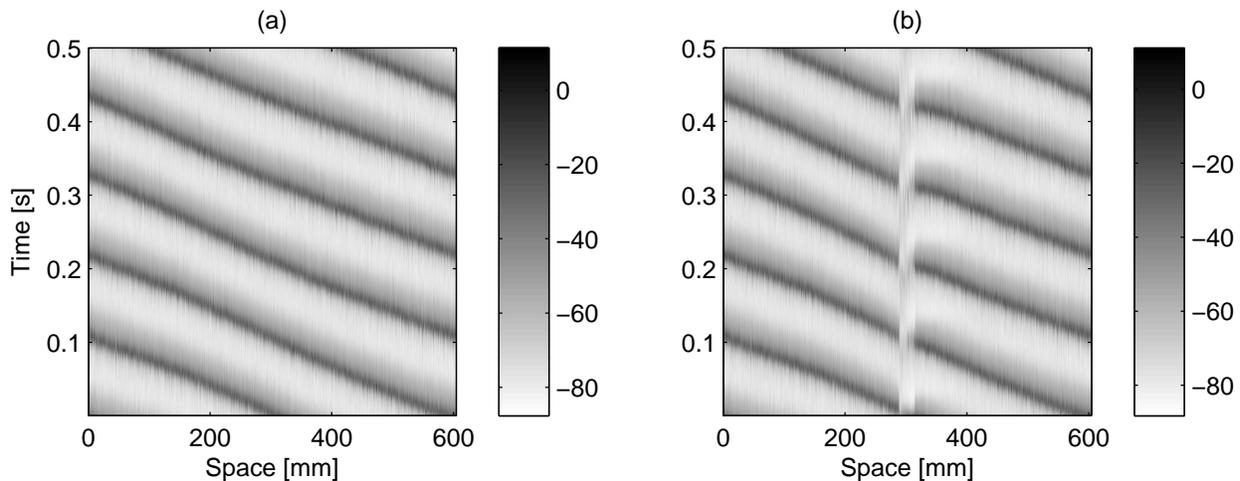


Figure 4.10: (a) Traveling waves in h_e for $\lambda_e = 11.2$, with no feedback control. This simulation used $\Gamma_e = 0.001$, $P_{ee} = 548.0$ and $\alpha = 5.0$. The waves travel from right to left with an approximate frequency of 9Hz and speed of 3m/s. (b) Traveling waves from (a), with the addition of one electrode performing proportional feedback control as described in (4.4). Here $a_{max} = 2$, $b = -0.1$, and the electrode width is 20.16 mm.

order of 15mV; when the control is turned on, this number will decrease. The amount that it decreases will depend on the gain of the controller and its ability to suppress the seizure waves.

For our analysis, we choose electrode sizes of 3.36mm, 10.08mm, and 20.16mm; the length of the hyperbolic tangent falloff is 5.6mm in every case. Before each simulation, we set the state of the random number generator in Matlab. This allows us to make direct comparisons between simulations because it assures that the cortical model receives the same stochastic input every time. As a baseline experiment, we run a simulation of uncontrolled cortex ($u = 0$) for each set of parameters.

First, we vary the size of the electrode at a constant value of λ_e . Figure 4.11 shows the the standard deviation at each point in space *minus* the uncontrolled standard deviation at that point for (a) $\lambda_e = 5.6$, (b) $\lambda_e = 11.2$, and (c) $\lambda_e = 22.4$. We can therefore think of these traces as the change in standard deviation due to the application of feedback control. After subtraction, the signals were smoothed with a 10-point moving average. Each subfigure shows the results for three electrode sizes.

For all three length scales, we see a “bump” in standard deviation at the front (right) edge of the electrode, analogous to the build-up of a physical wave as it meets resistance; recall that the electrical waves are traveling right to left. This increase in standard deviation is greater for larger electrodes. On the other hand, we see a decrease in standard deviation on the trailing (left) side of the electrode, and this decrease in wave amplitude is greater for longer electrodes. There are differences in standard deviation underneath the electrodes, as well. Figure 4.12 shows a close-up view of the results in Figure 4.11. These pictures demonstrate that the most effective control (i.e. the greatest decrease in standard deviation)

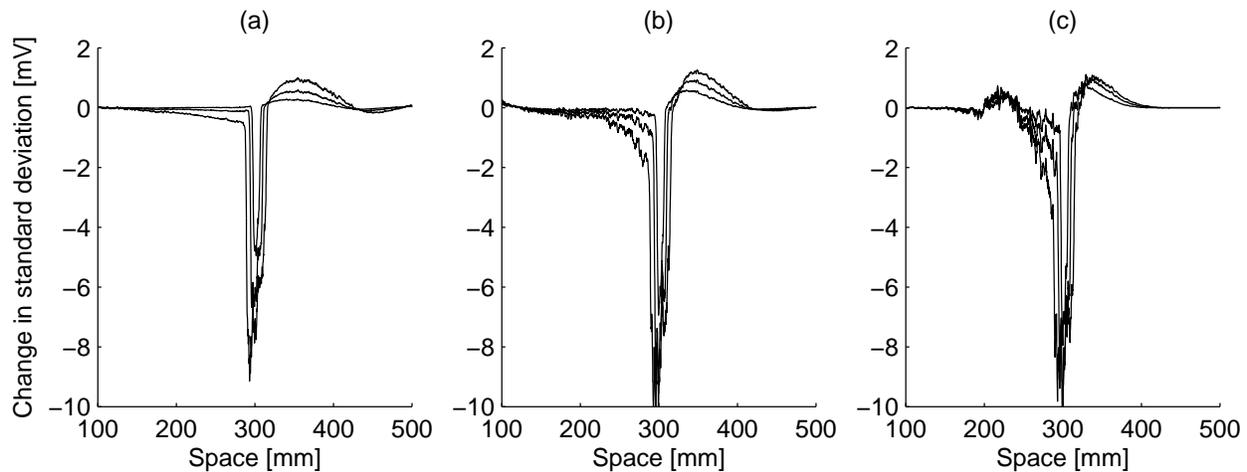


Figure 4.11: Change in standard deviation of h_e relative to uncontrolled seizure waves, for (a) $\lambda_e = 5.6$, (b) $\lambda_e = 11.2$ (typical value), and (c) $\lambda_e = 22.4$. After subtraction, the signals were smoothed with a 10-point moving average for ease of presentation. Each subfigure has traces for three different electrode sizes: 3.36mm, 10.08mm, and 20.16mm; the electrodes applied proportional control with $a_{\max} = 2$ and $b = -0.1$. Here we see that larger electrodes provide a greater reduction in standard deviation underneath the electrode. They also lead to a greater increase in standard deviation in front of the electrode and a greater decrease after the trailing edge. Note that the waves travel from right to left.

generally occurs on the trailing end of the electrode. We can also see that larger electrodes provide a greater decrease in standard deviation for the same control gain. Therefore, there is a tradeoff in the choice of electrode size. Larger electrodes have a greater build-up of electrical activity on their leading edge, but can provide better control underneath the electrode for the same control gain. Smaller electrodes generally have less effect on the surrounding tissue, but may require a higher gain for the same decrease in standard deviation.

Next, we compare the standard deviation measurements for a constant electrode size and a variation in λ_e . Figure 4.13 shows the results for an electrode of width 3.36mm and three different values of λ_e all plotted on the same axis; these traces were taken directly from Figure 4.11. In Subfigure 4.13(a), we see the results for the entire simulation range; a close-up view under the electrodes has been provided in Subfigure 4.13(b). Here we see that the spatial extent of the control is *independent* of λ_e underneath the electrode. Despite a quadrupling of the spatial length scale, the falloff of control at the electrode edges occurs at approximately the same position. This is also true for the activity leading up to the electrode; although the highest value of λ_e leads to the greatest increase in standard deviation, the return to uncontrolled levels happens at roughly 400mm for every value of λ_e . The only exception to this appears to be $\lambda_e = 22.4$ in the section of cortex following the electrode. The combination of parameters for that simulation results in an overshoot of the change in standard deviation before it returns to zero. Therefore, application of control with $\lambda_e = 22.4$ affects the traveling waves until a position of 200mm, whereas the waves for $\lambda_e = 5.6$ and 11.2 return to uncontrolled levels very quickly and with no overshoot.

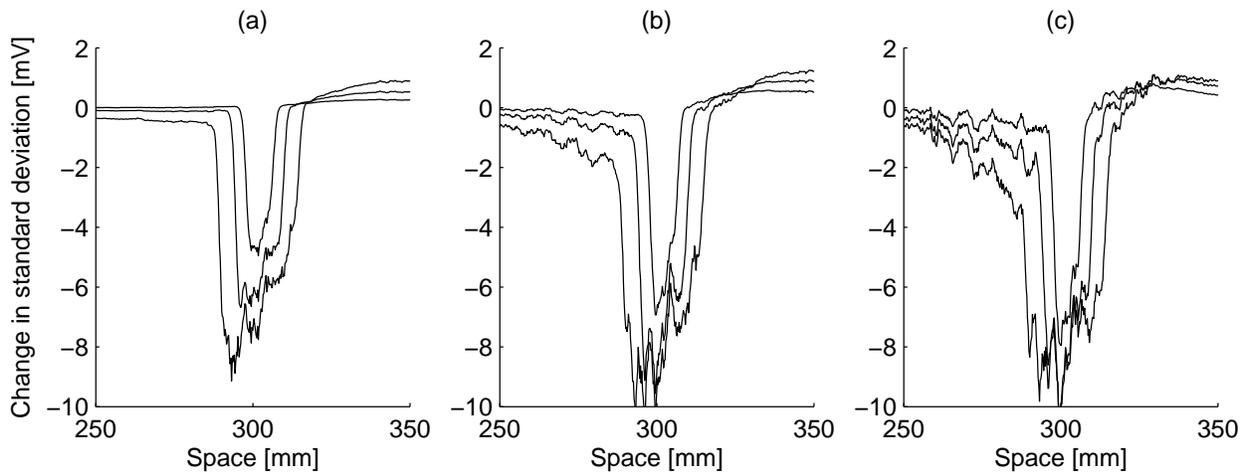


Figure 4.12: Close-up views of the results in Figure 4.11 showing the variation beneath the electrode. Note that the degree of control is greater at the trailing (left) edge of the electrode, as evidenced by the larger decrease in standard deviation. In addition, larger electrodes provide a greater overall decrease in standard deviation while using the same control gain.

These results appear to contradict the hypothesis stated at the beginning of this section, that the electrode size needed for successful control would be related to the length scale of the cortex. However, it seems that the electrode lengths feasible for feedback control are not big enough to have problems related to the cortical length scale. For example, at $\lambda_e = 11.2$, the seizure oscillations have a wave length of roughly 30cm. Any electrode over this length would sense more than one wave at a time and therefore provide inaccurate feedback. Because this electrode length is much too large for implementation, it is not a cause for concern.

Analysis of this kind can suggest strategies of electrode design for feedback control experiments. For example, a large electrode will require a low gain and therefore may be the safest option with regards to the cortical tissue. However, larger electrodes also lead to an increase in the magnitude of the electrical activity leading up to the controlled area, and this could potentially be a harmful effect. Moreover, a large electrode may span a sulcus or other structure, which may change the conclusions reached by exercising a homogeneous cortical model. Studies of these competing factors could produce a range of acceptable electrode sizes from which the experimentalists could choose. It would also be worthwhile to simulate the effect of two electrodes placed in close proximity to one another. Perhaps with this configuration, the second electrode could apply control to reduce the build-up of activity on the leading edge.

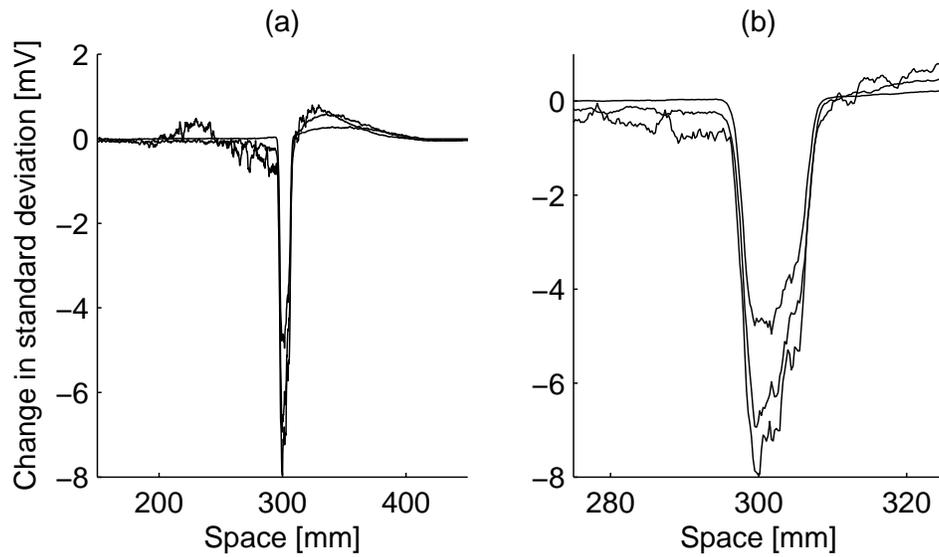


Figure 4.13: (a) Change in standard deviation for three different values of λ_e , all with an electrode size of 3.36mm. As in previous figures, $\lambda_e = 5.6, 11.2,$ and 22.4 . The data was taken directly from Figure 4.11. (b) A close-up view of the results in (a). These plots show that the spatial spread of the applied signal beyond the edges of the electrode is essentially independent of λ_e . The only exception to this is the trailing (left) edge of the electrode with $\lambda_e = 22.4$, which shows an overshoot in standard deviation before returning to uncontrolled levels. The pictures for other electrode sizes are qualitatively similar.

Chapter 5

A continuous mapping of EEG sleep states

The model of the sleep cycle presented in Section 2.2.2 provides an opportunity to explore human sleep stages in more detail. This refined characterization of states may provide insight into the point within the sleep cycle at which seizures are most likely to occur. For example, it could tell us not only that the subject is in stage 2 sleep, but also that the sleep is getting deeper and that a transition to stage 3 is near.

To obtain a continuous mapping of sleep states, we must compare the cortical model directly to human EEG recordings. This is done by using the sleep manifold as a way to generate model “EEG-like” signals. We will choose values of L and Δh_e^{rest} , find the numerical solution of the model for a given length of time, convert the dimensionless \tilde{h}_e to mV, and downsample it to match the EEG recordings. By doing this for many different values of L and Δh_e^{rest} we can obtain representative signals of every sleep stage.

We then calculate a set of features (Section 3.2.1) for both the EEG and model data and apply the LLE algorithm (Section 3.2.2) after identifying the largest strongly connected component (Section 3.2.3). For background and details on each of these concepts, please refer to the associated section.

5.1 Sleep EEG

5.1.1 Traditional EEG sleep stages

As was mentioned in Chapter 1, human sleep is traditionally divided into five categories: rapid eye movement (REM) sleep plus four stages of non-REM sleep (called stages 1-4). Stages 3 and 4 are often grouped together and called *slow-wave sleep* because they are dominated by low frequency delta waves. As a person moves from light sleep (stage 1) to deep sleep (stage 4), there is generally an increase in EEG amplitude and a decrease in the frequency. However, each stage has its own characteristics [18, 36]:

- **Awake:** The EEG of an awake person is characterized by low amplitude EEG signals, generally 10-30 μV and 16-25 Hz. However, this activity can often be obscured by motion artifacts.

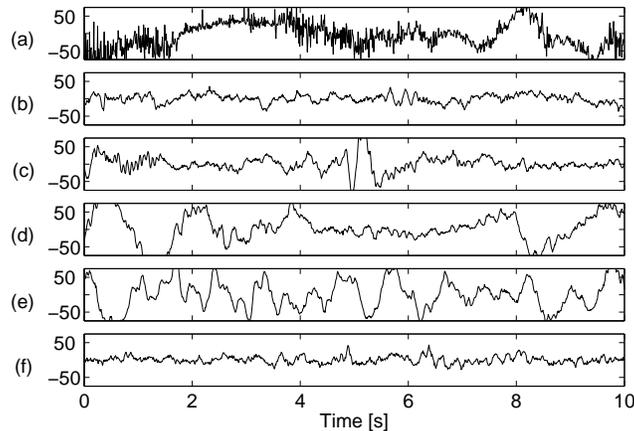


Figure 5.1: Examples of traditional EEG sleep stages, taken from PhysioBank data set sc4002e0: (a) awake, (b) Stage 1, (c) Stage 2, (d) Stage 3, (e) Stage 4, and (f) REM. Here, movement of the subject and EEG electrode appear to mask the typical awake behavior. As sleep deepens from stage 1 to stage 4, there is generally an increase in amplitude and a decrease in frequency. REM sleep, characterized by rapid eye movements, has a very low amplitude and resembles stage 1.

- **Stage 1:** The transitional stage between waking and sleeping is called stage 1; it is often thought of as the state of drowsiness. The EEG signal has a small amplitude and mixed-frequencies, and it is characterized by “alpha dropout,” a decrease in the alpha waves present when the subject is relaxing with closed eyes.
- **Stage 2:** Stage 2 sleep is defined by several unique components: bursts of activity in the 12-14 Hz range called *sleep spindles* and high amplitude biphasic waves named *K complexes*. The background activity is low in amplitude and in a lower frequency range.
- **Stage 3:** Compared to stage 2 sleep, stage 3 is higher in amplitude and contains periods of prominent delta waves in the range of 0.5-2 Hz. Sleep spindles may still be present in this stage, but are less prominent.
- **Stage 4:** Stage 4 is similar to stage 3, but it is more strongly dominated by low frequency delta activity. Spindles are rare in this stage.
- **REM:** As its name implies, REM sleep is characterized by rapid movement of the eyes, but otherwise looks very similar to stage 1 sleep. It has a low amplitude and a pattern of mixed frequencies. This is also the stage where most dreaming occurs.

A sample EEG trace for each of these stages is shown in Figure 5.1.

5.1.2 PhysioBank data sets

The EEG data used in the following analysis was obtained from the Sleep-EDF database [19], which is part of the PhysioBank online resource of physiologic signals for biomedical

research [14]. We used four data sets (sc4002e0, sc4012e0, sc4102e0, and sc4112e0), each one consisting of a European data format (EDF) file and a file containing the hypnogram data. They were converted to ASCII format and then imported into Matlab.

The data was gathered in 1989 from healthy Caucasian males and females between the ages of 21 and 35. Recordings were obtained over the course of one full day and include horizontal electrooculogram (EOG), two channels of EEG (Fpz-Cz and Pz-Oz sampled at 100 Hz), submental-electromyogram (EMG) envelope, oro-nasal airflow, and rectal body temperature. However, we used only the data from the Fpz-Cz EEG electrode pair in our analysis. The hypnogram data was generated via manual scoring according to Rechtschaffen & Kales using the two channels of EEG. For more details on the subjects, recording methods, and sleep staging, please see the full description in [34].

5.2 LLE applied to sleep EEG data

Before examining the connection between EEG data and the mathematical model of the sleep cycle, we will first discuss the results of applying LLE to sleep EEG only. After introducing the data sets and our methods, we show that LLE can separate EEG data by sleep stage and provide a continuous representation of sleep depth.

5.2.1 LLE input based on EEG features

In order to use the EEG as an input to the LLE function, we need to define our high-dimensional data set. We do this by dividing the signal into non-overlapping windows and calculating both statistical and frequency-based *features* for each one. Therefore each window becomes one high-dimensional data point, where the dimension equals the number of features. Because the data was scored using 30-second epochs, this was a natural choice for the window length. Thus, if we have 100 minutes of EEG data and we calculate six features, we will input 200 six-dimensional points into LLE and seek the embedding in two dimensions.

We start with a pool of 17 features and use various subsets to perform the LLE analysis; after the initial calculation of the features, each one was scaled by its root mean square (RMS) value. An algorithm for the automated selection of feature combinations is discussed in Section 5.2.2. The features are as follows:

- 1-5. Power in the delta (up to 4 Hz), theta (4-7.5 Hz), alpha (7.5-12 Hz), beta (12-26 Hz), and gamma ranges (above 26 Hz)
6. Total power
- 7-9. Variance, skewness, and kurtosis
10. Spindle score
11. Permutation entropy
- 12-13. Slope and offset of a linear fit to the power spectral density (PSD)
- 14-15. Peak height of the PSD above the linear estimate in the delta and alpha ranges
- 16-17. Fraction of power in the low (delta and theta) and high (beta and gamma) ranges

Full descriptions of these features are listed in Section 3.2.1.

The selection of a subset of features from this list may seem like a difficult task. It is certainly an important one – the use of all 17 features or a “nonsensical” subset will give poor results. However, it is worth noting that there are *many* combinations that result in a satisfactory separation between sleep stages in the LLE embedding. While each one may be slightly different, there will be a large number of high quality with respect to discrimination.

5.2.2 Automated ranking of feature sets

When we apply LLE to the EEG data, there are essentially only three choices that we must make:

1. *How many nearest neighbors should we include?* In other words, what is the value of k ? The LLE embedding will be stable over a range of values; we generally expect that k will be greater than the number of output dimensions and smaller than the original number of dimensions D [48].
2. *What should be the dimensionality of the LLE output space?* A nice property of the LLE algorithm is that each dimension is preserved as additional dimensions are added. Therefore, if we look at the results in two dimensions and do not achieve the desired mapping, we can add a third dimension without affecting the first two.
3. *Which combination of features should we use?* Employing all 17 features in our LLE analysis does not guarantee good separation between sleep stages because some of the features may not show consistent variation as the sleep depth changes. In addition, some features, such as the variance and the power in the delta band, show similar trends; we may achieve better results by eliminating these redundancies.

In this section, we focus on the last of these questions.

While we were able to identify many effective feature combinations through educated guesswork, we wanted to evaluate the utility of LLE as a method of sleep staging by identifying the best possible results. In this case, the “best” results are those that provide a large separation between sleep stages, especially between REM and deep slow-wave sleep. Because testing each combination of the features is an onerous task, e.g. choosing six features from a pool of 17 results in 12376 combinations, we developed an algorithm to evaluate the results automatically. It first identifies two groups of points: those marked as REM in the hypnogram and those determined to be stage 4. It then tracks two parameters based on the separation between those two groups of data points as they are embedded in the LLE output space.

We first measure the percent separation between REM and stage 4, calculated as

$$a_i = 100 \cdot \operatorname{erf} \left(\sqrt{0.5} \frac{\mu_4 - \mu_{REM}}{\sigma_4 + \sigma_{REM}} \right), \quad (5.1)$$

where μ and σ are the mean and standard deviation, respectively. This is based on the assumption that the best separation occurs when the distance between the means is large

and the total standard deviation is small. We perform this calculation in both the Y_1 and Y_2 directions and combine those measurements using the 2-norm to obtain the first parameter:

$$A = \sqrt{a_{Y_1}^2 + a_{Y_2}^2}. \quad (5.2)$$

The second parameter B uses the concept of nearest neighbors to evaluate separation; for example, if the stage 4 data points have only other stage 4 points as nearest neighbors, then we can infer that they are completely separated from the other sleep stages. More specifically, it measures the number of stage 4 points with REM points as nearest neighbors and divides that by the total number of stage 4 points. If the stage 4 group is isolated, we will have $B = 0$.

We determined the values of A and B for all possible combinations of six features. There were 267 feature sets where A exceeded a threshold of 90% separation in each direction: $A > \sqrt{90^2 + 90^2}$. We then identified the 267 feature sets with the lowest values of B . By finding the combinations that were common to both groups, we identified the 11 best feature sets. Visual inspection of the LLE results for these combinations confirmed the desired separation between REM and stage 4 sleep. Note that all 11 of these feature combinations provided results with $B = 0$.

5.2.3 Separation of sleep stages via LLE

Having described the EEG data set using frequency-based and statistical measures and having identified the most effective subsets of those features, we are now ready to apply the LLE algorithm. As a representative result, we choose one of the 11 feature sets from the previous section; the six features are power in the delta and theta bands, variance, spindle score, maximum height of the PSD above a linear estimate in the alpha band, and high power fraction. These are plotted in Figure 5.2 for 178 epochs from the sc4002e0 data set. The corresponding hypnogram is included for reference. Note that the features were calculated in 30-second non-overlapping windows to match the sleep scoring of the hypnogram.

We then use these features as the high-dimensional input to the LLE algorithm. The 2D results for 13 nearest neighbors ($k = 13$) are displayed in Figure 5.3(a). Every point in this figure represents a 30-second window of EEG data, and the color and symbol represent the sleep stage as determined by manual scoring. Here we see a very clear separation between the REM points (red circles) and those from stage 4 (blue stars), as required by our criteria for the automatic selection of the feature set. Stages 1 through 3 are located between those two groups and are arranged by sleep depth. In this example, we see a general trend of increasing sleep depth as we move to the upper right corner of the space. In addition, this low-dimensional embedding provides results with a smooth temporal progression. This is demonstrated by Subfigure 5.3(b), where the LLE results from Subfigure 5.3(a) are plotted versus time. In this example, the gradual transition to deep stage 4 sleep and the quick transition to REM are visible in the plot of Y_1 .

We would like to emphasize the importance of identifying strongly connected components when using LLE. Subfigure 5.4(a) shows an example of the Y_1 - Y_2 output space when LLE is performed on all 178 data points. Here, the feature set consisted of power in the delta, theta, and gamma bands, total power, maximum value of the PSD in the alpha band, and the low

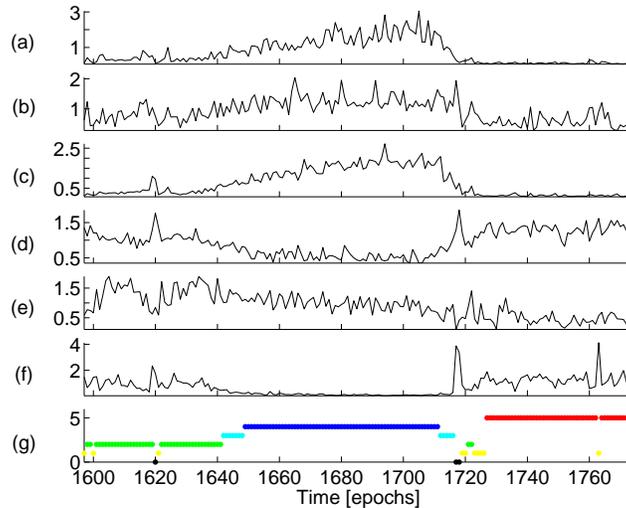


Figure 5.2: Scaled features of EEG data set sc4002e0, as described in Section 5.2.1. The subfigures show power in the (a) delta and (b) theta bands, (c) variance, (d) spindle score, (e) maximum height of the power spectrum in the alpha band after subtraction of a linear estimate, and (f) high power fraction. Subfigure (g) shows the hypnogram of the EEG data, where the number and color indicate the sleep stage: awake (0, black), stage 1 (1, yellow), stage 2 (2, green), stage 3 (3, cyan), stage 4 (4, blue), and REM (5, red). The features were calculated for the data from epochs 1597-1774 in 30-second windows with no overlap.

power fraction. Again, each point represents 30-seconds of EEG data, and the symbol (and color) are assigned based on its designated sleep stage. While there is some visible separation between the stages, the overall trend is unclear. On the other hand, Subfigure 5.4(b) shows the results when LLE is applied to the largest strongly connected component within the data. This component was identified as described in Section 3.2.3, and all other points were removed before using LLE. This greatly improves the results; the data points are spread further apart, and we see a grouping of sleep stages similar to Figure 5.3. Sometimes the removal of weakly connected points has a very small impact on the results, but situations like this make it a necessity. The significant improvement for this feature set allowed it to be counted as one of the “best” 11 results discovered by the automatic algorithm.

In this case, analysis of the strongly connected components resulted in the removal of eight data points:

- 3 points from waking (epochs 1620, 1717, and 1718)
- 2 points from stage 2 (epochs 1619 and 1624)
- 2 points from stage 3 (epochs 1712 and 1713)
- 1 point from stage 4 (epoch 1665)

Based on Subfigure 5.4(a), we can see why some of these were removed; there are four points that are clearly isolated from the rest of the data. However, the removal of points from

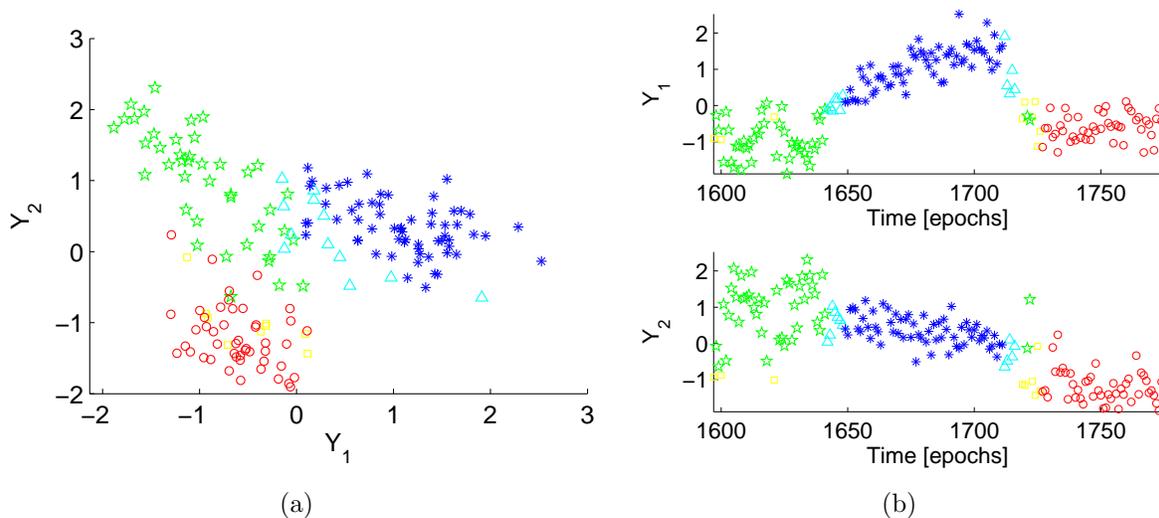


Figure 5.3: (a) Results of applying LLE to EEG data using the six features in Figure 5.2. The features were calculated for 30-second non-overlapping windows of data and the resulting 6-dimensional points were embedded in 2D space using LLE with $k = 13$; therefore, each point in this figure represents 30 seconds of EEG that has been characterized by the six features. The color and shape indicate sleep stage based on manual scoring: awake (black +), stage 1 (yellow \square), stage 2 (green \star), stage 3 (cyan \triangle), stage 4 (blue $*$), and REM (red \circ). (b) LLE output dimensions Y_1 and Y_2 versus time, for the results shown in (a). This demonstrates that LLE provides a low-dimensional output where the sleep state changes smoothly over time.

stages 3 and 4 are much less obvious. It is important to realize that, by using the concept of strongly connected components, this decision is automatic – it allows us to avoid the subjective selection of outlier points.

5.3 Integration of EEG data and the model sleep cycle

Thus far, we have shown that LLE is capable of distinguishing between sleep stages using only one channel of EEG and that the embedding exhibits a smooth progression over time. However, remember that our original goal was to find the relationship between EEG data and the mathematical model of the sleep cycle. Here we accomplish this by applying LLE *simultaneously* to EEG data and simulated data based on the model variable \tilde{h}_e .

It was previously argued that \tilde{h}_e cannot be directly compared to measurements from cortical surface or scalp electrodes because those measurements are based on extracellular current flow, as opposed to the soma potential. This is important for the modeling of certain cortical phenomena; for example, in performing simulations of feedback control for the suppression of epileptic seizures, the value of the electrode measurement is fed directly back to the cortex to affect \tilde{h}_e , with little or no time delay [30]. In that case, the relationship between \tilde{h}_e and the electrode measurement at any given time is very important. However, in the present analysis of EEG data using LLE, we are only interested in matching *scaled*

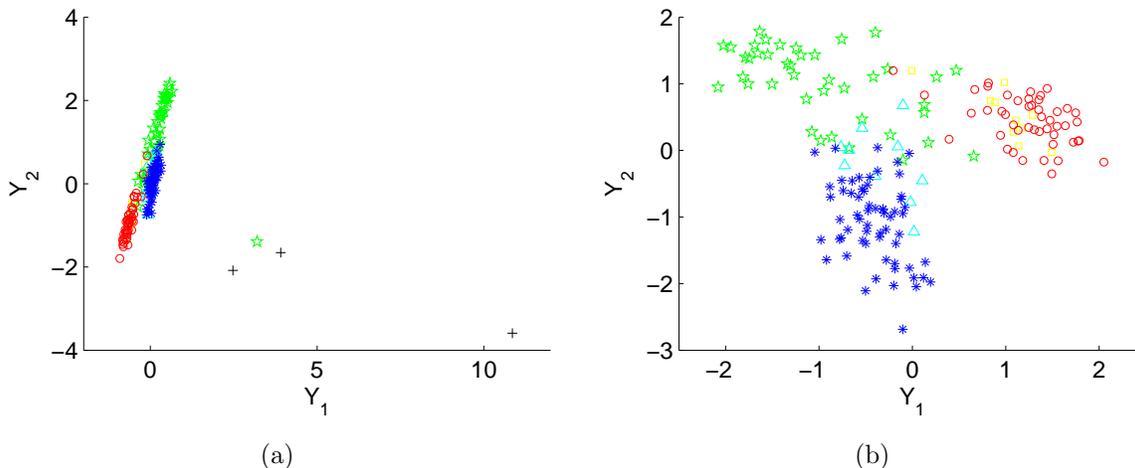


Figure 5.4: LLE results on sleep EEG data before (a) and after (b) removal of eight weakly connected points. The six features were power in the delta, theta, and gamma bands, total power, maximum height of PSD above a linear estimate in the alpha band, and low power fraction, and we used $k = 13$. As before, the color and symbol indicate sleep stage: awake (black +), stage 1 (yellow \square), stage 2 (green \star), stage 3 (cyan \triangle), stage 4 (blue $*$), and REM (red \circ). Note the dramatic improvement in separation between sleep stages when LLE is done on only one strongly connected component in (b).

features of the data that are calculated over 30-second intervals. We will not attempt to compare the temporal progression of \tilde{h}_e directly to the EEG data. The previous work mentioned above has demonstrated a correspondence between \tilde{h}_e and sleep EEG data with regard to these general features, so we feel confident in using it in the following analysis without the addition of a scalp electrode model.

A remark about principal component analysis (PCA) is in order. This is perhaps the most common mode of dimensionality reduction, and it has also been used in the analysis of sleep EEG data [12, 17, 6]. However, PCA places the greatest importance on the directions of largest variance and relies on the assumption that the data is best reconstructed by a *linear* combination of the original measurements. While we tried PCA and achieved reasonable results, the nonlinear nature of the sleep manifold suggests that a more sophisticated solution is necessary. In addition, the concept of nearest neighbors on which the LLE algorithm is based enabled improvement in the separation of different sleep stages (see Section 5.2.3), and it will play a crucial role in defining the quantitative relationship between the EEG data and mathematical model, as is discussed in Section 5.3.2.

5.3.1 Model data set

To generate the model data set, we place a grid of points on the sleep manifold (Figure 2.4) and obtain the numerical solution of the cortical model at each one. We vary L over the interval $[0.5, 2]$ in increments of 0.1 and Δh_e^{rest} over $[-5, 5]$ in increments of 0.5. This gives us a total of 336 model signals for analysis; we then remove the initial transients and characterize each signal based on a subset of the features described in Section 5.2.1. In this

way, the nonlinear sleep manifold is turned into “EEG-like” signals which are converted to high-dimensional data points for use with LLE.

For the model data set, the length of each signal is 10 seconds (as opposed to the 30-second windows used for the EEG data). We are able to use this shorter time because we can choose parameters in the model to simulate a stationary brain state, i.e. we can use constant values of L and Δh_e^{rest} . A test of the feature calculations for various window lengths indicated that, in many cases, the signal properties were stationary for windows greater than five seconds. Certain parts of the sleep manifold had transients lasting roughly 10 seconds.

In order to compare this model data set directly to EEG measurements, it is important that all of the basic properties match. For example, just as REM EEG signals have a much lower variance than those from stages 3 and 4, we expect that the signals from the topmost REM portion of the sleep manifold will have a smaller variance than those on the lower NREM section. However, we found that the use of a constant α , which defines the variance of the stochastic input to the model cortex in equation (2.10), does not reproduce this behavior. Therefore, we varied the value of α as we moved in the L - Δh_e^{rest} space. More specifically, we based it on the sleep manifold. Define $\vec{\mu}_e$ to be a matrix of the steady-state values of h_e after they have been shifted and scaled to have a range of $[0, 1]$. Then we define a matrix of α values:

$$\vec{\alpha} = \alpha \cdot (-7\vec{\mu}_e + 8) . \quad (5.3)$$

Therefore, the REM portion of the model sleep cycle (where $\mu_e \sim 1$) will have stochastic inputs of α , while the lower NREM section (where $\mu_e \sim 0$) will have inputs of variance 8α . This stochastic input allowed us to successfully reproduce the desired range of variances in the model signals.

In addition to the variance, other features of the model data set mimic characteristics of sleep EEG. This can be verified by plotting the features as we traverse the sleep manifold. For example, power in the delta band, composite permutation entropy index (CPEI), “peak” height of the power spectral density in the alpha band, low power fraction, and high power fraction are shown in Figure 5.5. The values of each feature are displayed for the grid of points in L and Δh_e^{rest} that covers the sleep manifold. For reference, the steady state values of h_e on the sleep manifold are shown in Subfigure 5.5(a); note that this is similar to viewing Figure 2.4 from the top and coloring the points based on their height. The lowest value is plotted in white and the highest value in black.

As desired, Subfigure 5.5(b) indicates that the power in the delta band increases as the depth of sleep increases, with the largest values occurring near the quick transition to REM sleep. Similarly, Subfigures 5.5(e) and (f) show that the fraction of power in the low frequencies is greater during NREM sleep, while the fraction of power at high frequencies is greater during REM sleep. Consistent with previous reports that the CPEI decreases with depth of anesthesia [39], we see in Subfigure 5.5(c) that the CPEI decreases with sleep depth in the model. Subfigure 5.5(d) shows that the region of greatest alpha power is located in the upper left corner, for small values of L and large values of Δh_e^{rest} . As a means of comparison, the same five features were applied to a sample of EEG data and are displayed in Figure 5.6.

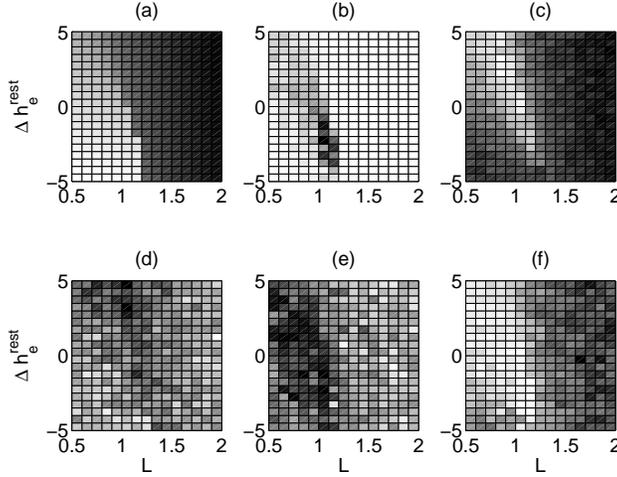


Figure 5.5: Variation of five features as the surface of the sleep manifold is traversed in $L-\Delta h_e^{rest}$ space. Each feature has been scaled by its RMS value and depicted in grayscale, with white indicating the lowest values and black representing the highest values. (a) The steady state values of h_e from the sleep manifold in Figure 2.4. The black points represent the upper REM branch, the white points represent NREM, and the fold is located at roughly $L = 1.2$. The other subfigures show (b) power in the delta band, (c) permutation entropy, (d) maximum height of PSD above a linear estimate in the alpha band, (e) low power fraction, and (f) high power fraction. These five features use α as defined in (5.3). They show that the representation of REM and NREM in the model is consistent with the characteristics of sleep EEG.

5.3.2 Application of LLE to a hybrid data set

We now join the EEG measurements and the model data into one hybrid data set and use it as an input to the LLE algorithm. This simultaneously finds the low-dimensional embedding for both data types and allows us to infer a correspondence between them. For example, Figure 5.7(a) shows the result of applying LLE to the grid of 336 model points and a full night’s sleep from EEG data set sc4002e0 (epochs 800 to 2000). The input data was composed of the five features from Figure 5.5: power in the delta band, CPEI, maximum height of the PSD in the alpha band (relative to a linear estimate), and the low and high power fractions. We used $k = 14$, and only three points were removed by analysis of the strongly connected components.

In Figure 5.7(a), the model data is represented by dots, where the color denotes the steady-state value of h_e associated with that point; in general, the red points represent the REM portion of the manifold, while the blue points represent NREM. On the other hand, the sleep EEG data points are rings, where the color is chosen based on sleep stage. Note that, for clarity, only the first 500 EEG data points were included in the figure.

The most important aspect of this result is that the EEG data points and model points overlap each other in the Y_1 - Y_2 output space. This implies that model points have EEG data points as nearest neighbors (and vice versa) and verifies that LLE has associated the two data

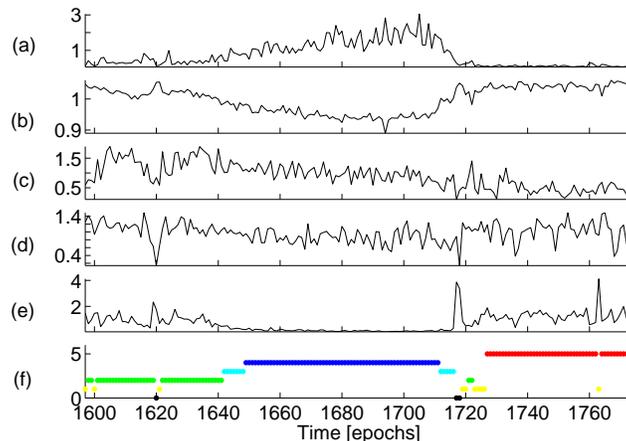


Figure 5.6: Variation of the features from Figure 5.5 when they are applied to sleep EEG data, rather than model data. The sample of EEG data was taken from sc4002e0, and each feature has been scaled by its RMS value. The subfigures show (a) power in the delta band, (b) permutation entropy, (c) maximum height of PSD above a linear estimate in the alpha band, (d) low power fraction, (e) high power fraction, and (f) hypnogram of the EEG data. The colors and numbering for the hypnogram are the same as those used for Figure 5.2. Note that the values of these features (relative to sleep stage) are consistent with the model results in Figure 5.5.

types with one another. Without fidelity of the model and careful choice of EEG features, we would have likely obtained a result with one cluster of EEG points and a completely separate cluster of model points. Further, LLE appears to have matched the sleep stages between the two data types – the deepest sleep (blue for both EEG and model) appears in the lower left corner, and REM (red) is embedded in a vertical band where Y_1 is in the range $[-1, 0]$. The separation between sleep stages can be seen more clearly in Subfigure 5.7(b), which displays only the EEG data points from Subfigure 5.7(a). Here we see that the stages are grouped; even the REM points and the awake points are separated, despite the fact that their EEG traces are characterized by very similar features. If we were to plot the Y_1 and Y_2 values of the EEG data points as they evolve in time, we would see a very similar result to the one in Figure 5.3(b). Here, the Y_1 direction appears to be an approximate indicator of sleep depth.

5.3.3 Connection to the theoretical sleep manifold

So far, we have seen that LLE provides a qualitatively similar embedding for REM and NREM points in both EEG measurements and simulated model data. However, we would like to quantify this relationship. In other words, we would like to associate each EEG data point with a position on the sleep manifold in the $L-\Delta h_e^{rest}$ space. This will allow us to infer the model trajectory of a subject’s actual brain state as it moves along the manifold.

To do this, we use the results in Figure 5.7(a) and again turn to the concept of nearest neighbors. Using $k = 14$, we calculate the nearest neighbors of every point in the Y_1 - Y_2

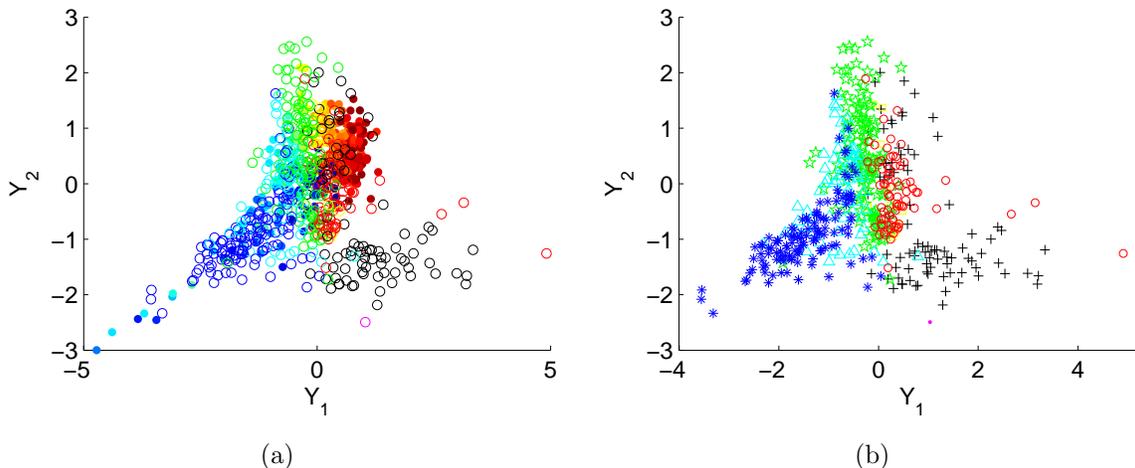


Figure 5.7: (a) LLE results for a hybrid data set containing both sleep EEG data and numerical solutions of the cortical model. We used the five features from Figure 5.5 and set $k = 14$. The rings represent EEG data and are colored by sleep stage. While the analysis included 1200 windows of EEG data, only 500 are displayed here for clarity. The solid dots represent data from the model; they are colored based on the mean value of h_e at that point, where red represents the highest (REM) values, and dark blue marks the lowest (NREM) values. Note that the data and model points overlap in the output space and that the arrangement of sleep stages is very similar. (b) LLE results showing the EEG data *only*, using the same colors and symbols as Figure 5.4. This allows us to see that the data has been roughly separated by sleep stage.

space. We then identify *model* points that are nearest neighbors of *EEG* data points. Each one of those model points has an associated position on the sleep manifold; we assume that the $L-\Delta h_e^{rest}$ positions of the model nearest neighbors will be the most closely associated positions for the EEG data point.

We can visualize this concept by creating histograms of the model nearest neighbors and separating them by sleep stage (Figure 5.8(a)). Every time a model point is a nearest neighbor of an EEG point, we increment the count at the model point's associated location in $L-\Delta h_e^{rest}$ for the sleep stage of the EEG point. We then create grayscale plots of the total counts, where white indicates that a location was never a nearest neighbor of that sleep stage and black indicates that it was a nearest neighbor many times.

For example, (i)-(vi) in Subfigure 5.8(a) correspond to awake, REM, and stages 1-4, respectively. The thick vertical line at $L = 1.2$ marks the approximate location of the fold. As we move from REM to the deeper stages of sleep, we can see a continuous progression along the sleep manifold. In this example, REM and stage 1 sleep generally associate themselves with locations on the right half of the manifold (and a small piece of the lower left corner). Then in stage 2 sleep, we move to the left half of the manifold; here, we see two distinct groups of points, with a majority landing in the group that borders the area associated with REM and Stage 1. Stage 3 is associated with a cluster of points starting in the upper left-hand corner and approaching the fold. Stage 4 continues this progression and is located in a band of points leading up to the fold.

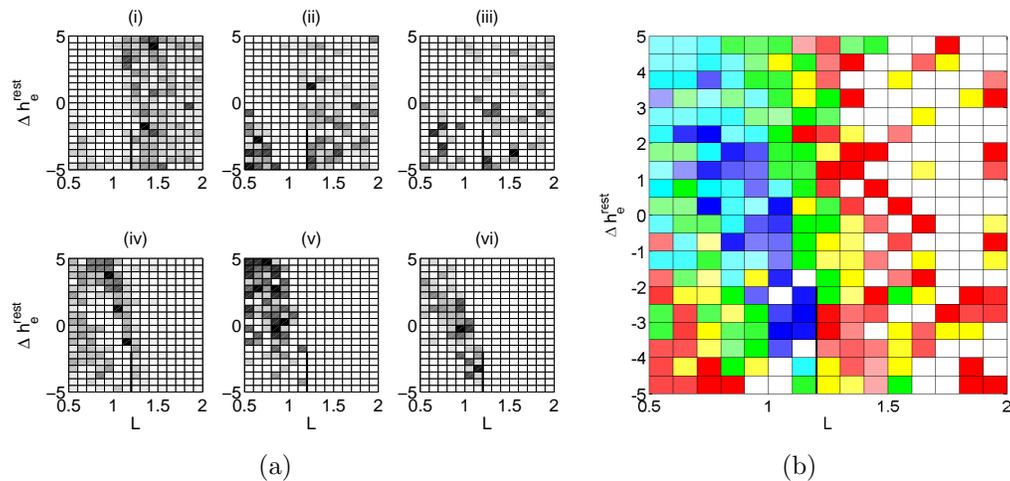


Figure 5.8: Association between EEG data set sc4002e0 and the sleep manifold. Each picture shows the sleep manifold in $L-\Delta h_e^{rest}$ space, with a heavy black line to indicate the location of the fold. (a) Histograms of nearest neighbors for (i) waking, (ii) REM, (iii) stage 1, (iv) stage 2, (v) stage 3, and (vi) stage 4 sleep. The shading of each square indicates the number of times that location on the sleep manifold was a nearest neighbor of EEG data in that stage. For example, (vi) shows that stage 4 sleep most often associates itself with the lower NREM branch of solutions leading up to the fold. (b) A composite picture of the results in (a), where each location is colored based on the sleep stage with the most neighbors at that point, relative to the total number of neighbors associated with that stage. Again, we use stage 1 (yellow), stage 2 (green), stage 3 (cyan), stage 4 (blue), and REM (red). The intensity of color is scaled based on the percentage of neighbors that come from that stage; the more saturated the color, the greater the percentage. Waking points were excluded.

We can then create a composite plot that combines all five sleep stages (we neglect the waking points for this task because the awake state was not included in the sleep cycle model). For every location on the manifold, we determine which sleep stage it was most closely associated with and color it accordingly. To do this, we scale the number of nearest neighbors for each stage by the total number for that stage; then, for every position on the manifold, we choose the stage with the highest value. This accounts for the fact that the subjects do not spend an equal amount of time in each sleep stage (otherwise, more time spent in a certain stage would lead to more nearest neighbors and a greater likelihood of dominating this composite plot). As in previous figures, we use red for REM, yellow for stage 1, green for stage 2, cyan for stage 3, and blue for stage 4. The intensity of the color is assigned based on the percentage of times it was associated with that sleep stage. Suppose a certain point on the manifold was a neighbor of stage 2 twelve times, a neighbor of stage 1 five times, and a neighbor of REM three times. We would color that point green to indicate stage 2 sleep, and its saturation value would be $12/(12 + 5 + 3) = 0.6$. In other words, the intensity of the color is a “confidence” measure; the more saturated the color, the more closely it is associated with that sleep stage. The composite figure for the data in Subfigure 5.8(a) is shown in Subfigure 5.8(b).

5.3.4 Inclusion of additional data sets

It is important that this method of analysis works consistently for different subjects with a variety of sleeping patterns. We tested this capability using the full night of sleep from each of the remaining three data sets: sc4012e0, sc4102e0, and sc4112e0. Rather than start from scratch and re-run the LLE algorithm, we projected the new data onto the existing embedding. For a new input \vec{x} , this is a three-step process [48]:

1. Find the k nearest neighbors of each new data point among the points in the existing embedding.
2. Compute the best linear reconstruction w_j of each new point using only its nearest neighbors. Again, we enforce the constraint that the weights used in the reconstruction sum to one: $\sum_j w_j = 1$.
3. Calculate the output for the new data points: $\vec{y} = \sum_j w_j \vec{Y}_j$, where \vec{Y} contains the original embedding coordinates and j cycles through the neighbors of \vec{x} .

This is more computationally efficient than running the entire algorithm again, and it guarantees that the output embedding will not change as we add new data. Most importantly, this makes it possible to do continuous real-time monitoring of EEG data; a new point could be projected onto the results every 30 seconds (or less) as the subject sleeps.

When we project the sleep data from files sc4012e0, sc4102e0, and sc4112e0 onto the embedding derived from sc4002e0, we obtain the composite pictures in Subfigures 5.9(a)-(c), respectively. All three results are consistent with one another, despite coming from different subjects and containing a minimal amount of stage 3 and 4 deep sleep. The only exception to this is stage 4 sleep in Subfigure 5.9(c); however, it is important to note that only 21 points out of 1100 were denoted as stage 4 sleep for this subject, and those points were not all consecutive. Therefore, the subject had only transient movements into stage 4 from stage 3, and it is perhaps not surprising that the results show the stage 4 EEG points mixed in with those from stage 3. Also note that the placement of the sleep stages in Figure 5.9 is consistent with the results in Figure 5.8.

Lastly, we combine the results from all four data sets (the original embedding with sc4002e0 plus three projected data sets) to produce Figure 5.10. The histograms in Subfigure 5.10(a) were created by a simple summation of the nearest neighbor histograms for all four data sets. The composite plot in Subfigure 5.10(b) was then generated according to the logic described in Section 5.3.3 using the combined histogram data. In all, these results are based on almost 40 hours of EEG data from four different subjects. Again, they are consistent with the individual results and they show a clear picture of the sleep manifold regions associated with each sleep stage. It is also noteworthy that only a handful of points on the sleep manifold (colored white in the composite picture) were never nearest neighbors of an EEG data point.

This picture may be very useful in the analysis of seizures during sleep. Imagine taking another new sleep EEG data set, this time from an epileptic subject, and projecting it onto these results. By following the location in $L-\Delta h_e^{rest}$ as the subject sleeps, we can get an idea of the sleep stage as it is traditionally defined, and *we can also identify that stage in*

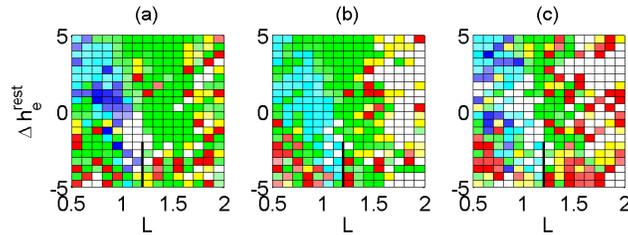


Figure 5.9: Composite plots for EEG data sets (a) sc4012e0, (b) sc4102e0, and (c) sc4112e0, when they are projected onto the LLE results from Figure 5.4(b), as described in Section 5.3.4. These pictures are analogous to Figure 5.8(b) and use the same color scheme. Note that the results are consistent with those for sc4002e0 in Figure 5.8(b); over various subjects, the sleep stages are generally associated with the same positions on the sleep manifold.

more detail and detect nearness to transitions between stages. The grid of points on the sleep manifold essentially gives us descriptions of 336 different brain states associated with sleep. We expect that future research will identify the locations on the sleep manifold where seizures are most likely to occur. With that knowledge, if the sleep state characterization is done continuously while the subject is sleeping, this may allow for the prediction (and possibly prevention) of seizures.

We emphasize the fact that the coloring in the composite pictures (Figures 5.8(b), 5.9, and 5.10(b)) is based on the subjective scoring of sleep data. It is well known that interscorer variability can be quite high, and this certainly affects our results. For example, imagine if some of the points scored as REM that landed in the range $1.6 < L < 2$ on the sleep manifold were instead scored as stage 1. Then the right side of the composite picture would be completely yellow and the region associated with stage 1 would be more clear. Therefore, the composite pictures should be seen as “guides” to tie the analysis back to the traditional definitions of the sleep stages, not as the ultimate truth. As mentioned in the previous paragraph, we are most interested in the *position* on the sleep manifold, the trajectory that results as the subject sleeps, and the relationship of this trajectory to the regions where seizures may be most likely to occur.

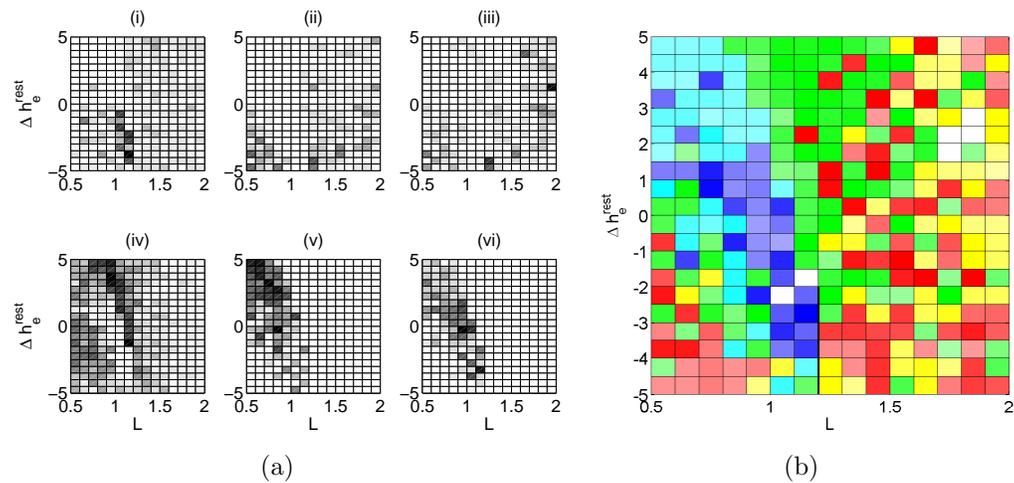


Figure 5.10: Combined association between all four EEG data sets and the sleep manifold. The data set sc4002e0 was directly compared to the cortical model using LLE, and the remaining three data sets were projected onto those results as described in Section 5.3.4. (a) Total histograms of nearest neighbors, separated by sleep stage; these were calculated by summing the histograms from all four EEG data sets. The pictures show awake, REM, and stages 1-4 in (i) through (vi), respectively. (b) The total composite picture for all four data sets. This was generated from the histogram data in (a) and is analogous to Figures 5.8(b) and 5.9. Again, this is consistent with previous results and shows the regions of the sleep manifold most closely associated with each sleep stage.

Chapter 6

Discussion

6.1 Summary of contributions

Mathematical models represent an opportunity for exploration and prediction. In this case, a mean-field, mesoscale model of the human cortex allows investigation into epileptic seizures, sleep, and the connection between them. The ability of this model to mimic the seizure waves seen in humans lets us simulate suppression via feedback control and make experimental recommendations. The representation of the sleep cycle contained within this cortical model creates the possibility for a more detailed description of sleep states, with application to the analysis and prediction of seizures during sleep.

First, we presented several novel approaches to the modeling of feedback control for epileptic seizures in humans. We verified the strong convergence of numerical solutions to the model of the cortex, paying special attention to discontinuities that may occur at electrode edges. This allowed us to choose appropriate step sizes for our simulations; because the spatial step size Δx was small relative to the size of the electrode, we were able to incorporate a more detailed electrode profile into the simulation. Then, based on evidence that the mean soma potential \tilde{h}_e cannot be used as the measurement for feedback control, we developed a new model \tilde{h}_m to represent the measurement of cortical surface electrodes. This model was based on the currents flowing in the cortex and was used for all simulations of feedback control. Those simulations utilized a new control algorithm containing the total integral of the applied potential u . Not only did this succeed in suppressing the seizure-like oscillations, but it guaranteed that the applied signal would be charge-balanced and therefore safe for cortical tissue.

Next, we used locally linear embedding to directly associate human sleep EEG data with the same mesoscale cortical model. We first showed that LLE has the ability to distinguish between sleep stages when applied to EEG data alone. This analysis can reliably separate REM and NREM sleep data and provide a smooth temporal progression through the various stages of sleep. We also presented the concept of strongly connected components as a method of automatic outlier rejection for EEG data and discussed a method for the selection of EEG features used in the analysis. Then, by using LLE on a hybrid data set containing both sleep EEG and signals generated from the mathematical sleep cycle, we were able to quantify the relationship between the model and the data. This enabled us to take any sample

of sleep EEG data and associate it with a position among the continuous range of sleep states provided by the model. In addition, this approach yields consistent results for various subjects over a full night of sleep and can be done in real time as the subject sleeps. This suggests a wide range of possibilities for future investigation, especially with regards to the prediction of seizures during sleep.

6.2 Future Work

While we have made several significant steps toward comparison of our seizure control model with experimental data, there is still much work to be done. In the analysis presented here, we have assumed that each electrode can be simultaneously sensing and applying the control signal. This is not realistic; ideally, we would model separate electrodes for these two tasks, with the geometric properties chosen to match existing experimental setups. Also, as mentioned previously, it would be possible for us to improve the electrode profile used in simulation. Here, we incorporated a simple profile to demonstrate our capability to do so, but it would be more accurate to base our choice on existing theories of the potential difference across an electrode surface [47]. It may also be possible to account for electrochemical changes that occur in the vicinity. Lastly, we note that our use of the phrase “charge-balanced” should be taken lightly. Our controller measures and applies a *voltage*, so the integral term pushes the total voltage (over time) towards zero. Although, in concept, this is similar to having a charge-balanced signal, it does not guarantee that the applied charges will be balanced and safe. This could be remedied by utilizing a controller that measures cortical potential and applies a *current*. Not only would this be more accurate, but it would facilitate future comparisons with experiments, most of which are done in this manner [5, 58].

Validation via experimentation is only one of the many possible future directions of this work. For example, we could extend our model to two dimensions and use it to study seizure waves, or we could simulate experimental phenomena such as irregular, spiral, and plane cortical waves [49]. Related theoretical work has suggested that pre-processing of data using a Kalman filter can provide greater flexibility in the control of waves while minimizing the amount of energy needed to do so [50]. This concept could be readily applied to the simulations discussed here. Another possible avenue of future work is the further investigation of spatial properties of our feedback model. Here we have analyzed the performance of single electrodes applying proportional control. However, we have the capability to do simulations with any number of electrodes at any size and spacing, with any type of control. This is a luxury not afforded to experimentalists. Continuing theoretical work in this area may provide insight into questions such as: if two electrodes are used in tandem, is it possible to prevent the build-up of activity that occurs as the seizure wave approaches? What is the best control strategy and spacing distance for this task? Can we achieve better results by using a different control strategy for each electrode? Where should these electrodes be placed, relative to the seizure focus, for maximum effectiveness? In general, what are the necessary resolutions for sensing and actuation?

Our analysis of sleep states in the EEG suggests many future projects, as well. The most direct path would be to use the same LLE analysis on sleep EEG data from epileptic subjects. By following the trajectory of their brain state on the sleep manifold during the

preictal period, perhaps a pattern of seizure onset could be discerned. It may be that seizures always occur in the same location on the manifold or that the trajectory leading up to the seizure has a characteristic trait. The analysis could be improved in several ways. First, it may be beneficial to use electrocorticogram (ECoG) data from the surface of the cortex, rather than EEG data. While there are no predefined notions of sleep stages for this data type, that is not a requirement for the algorithm presented here. However, it would be crucial to choose a set of features that accurately capture the characteristics of the ECoG; because it comes from under the skull (which acts as a low-pass filter), ECoG signals have larger amplitudes and contain more high frequency information. Second, it may be possible to devise a form of “fuzzy” LLE. Our analysis utilized a stochastic cortical model, but we took single numerical solutions and used them as representative of positions on the sleep manifold. We may obtain a more accurate mapping of EEG data to the model if we incorporated a *sampling* of numerical solutions from each point in $L-\Delta h_e^{rest}$ space.

Perhaps this work will eventually lead to the feedback control of epileptic seizures during wakefulness *or* during sleep, thus linking the ideas from Chapters 4 and 5. The framework discussed here would allow us to study the electrode geometry and control methods necessary for seizures that arise out of sleep. We could examine the effect of different types of control on the subject’s sleep state relative to the mathematical model; maybe it would be possible to steer the sleeping subject away from a dangerous state where seizures are likely to occur. It is our hope that this work will act as a stepping stone to such intriguing investigations.

Bibliography

- [1] M. Abeles. *Corticonics: Neural Circuits of the Cerebral Cortex*. Cambridge University Press, 1991.
- [2] P. Ataei, A. Yazdani, S. K. Setarehdan, and H. A. Noubari. Manifold learning applied on EEG signal of the epileptic patients for detection of normal and pre-seizure states. *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, pages 5489–5492, August 2007.
- [3] V. Braitenberg and A. Schüz. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, 2nd edition, 1998.
- [4] N. T. Carnevale and M. Hines. *Computational Neuroscience in Epilepsy*, chapter 2: The Neuron Simulation Environment in Epilepsy Research, pages 18–33. Academic Press, 2008.
- [5] M. E. Colpan, Y. Li, J. Dwyer, and D. J. Mogul. Proportional feedback stimulation for seizure control in rats. *Epilepsia*, 48(8):1594–1603, 2007.
- [6] M. Corsi-Cabrera, M. A. Guevara, Y. D. Río-Portilla, C. Arce, and Y. Villanueva-Hernández. EEG bands during wakefulness, slow-wave and paradoxical sleep as a result of principal component analysis in man. *SLEEP*, 23(6), 2000.
- [7] P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, 2001.
- [8] N. Foldvary-Schaefer and M. Grigg-Damberger. Sleep and epilepsy: What we know, don't know, and need to know. *Journal of Clinical Neurophysiology*, 23(1), February 2006.
- [9] G. F. Franklin, J. D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamics Systems*. Prentice Hall, 4th edition, 2002.
- [10] J. G. Gaines. *Numerical Experiments with S(P)DE's*, chapter in Stochastic Partial Differential Equations, pages 55–71. London Mathematical Society Lecture Note Series 216. Cambridge University Press, 1995.
- [11] J. G. Gaines and T. J. Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM Journal on Applied Mathematics*, 57(5):1455–1484, October 1997.

- [12] D. Gervasoni, S.-C. Lin, S. Ribeiro, E. S. Soares, J. Pantoja, and M. A. Nicolelis. Global forebrain dynamics predict rat behavioral states and their transitions. *The Journal of Neuroscience*, 24(49):11137–11147, December 2004.
- [13] B. J. Gluckman, H. Nguyen, S. L. Weinstein, and S. J. Schiff. Adaptive electric field control of epileptic seizures. *Journal of Neuroscience*, 21(2):590–600, January 2001.
- [14] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000 (June 13). Circulation Electronic Pages: <http://circ.ahajournals.org/cgi/content/full/101/23/e215>.
- [15] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546, 2001.
- [16] W. A. Hofstra, B. E. Grootemarsink, R. Dieker, J. van der Palen, and A. W. de Weerd. Temporal distribution of clinical seizures over the 24-h day: A retrospective observational study in a tertiary epilepsy clinic. *Epilepsia*, 50(9):2019–2026, 2009.
- [17] M. Jobert, H. Escola, E. Poiseau, and P. Gaillard. Automatic analysis of sleep using two parameters based on principal component analysis of electroencephalography spectral data. *Biological Cybernetics*, 71:197–207, 1994.
- [18] E. R. Kandel, J. Schwartz, and T. Jessell. *Principles of Neural Science*. McGraw-Hill, 4th edition, 2000.
- [19] B. Kemp. *The Sleep-EDF Database*. World Wide Web, <http://www.physionet.org/physiobank/database/sleep-edf/>, accessed August 2009.
- [20] M. Kinoshita, A. Ikeda, M. Matsushashi, R. Matsumoto, T. Hitomi, T. Begum, K. Usui, M. Takayama, N. Mikuni, S. Miyamoto, N. Hashimoto, and H. Shibasaki. Electric cortical stimulation suppresses epileptic and background activities in neocortical epilepsy and mesial temporal lobe epilepsy. *Clinical Neurophysiology*, 116:1291–1299, 2005.
- [21] M. A. Kramer, H. E. Kirsch, and A. J. Szeri. Pathological pattern formation and cortical propagation of epileptic seizures. *J. R. Soc. Interface*, 2:113–127, 2005.
- [22] M. A. Kramer, B. A. Lopour, H. E. Kirsch, and A. J. Szeri. Bifurcation control of a seizing human cortex. *Physical Review E*, 73:041928, 2006.
- [23] M. A. Kramer, A. J. Szeri, J. W. Sleight, and H. E. Kirsch. Mechanisms of seizure propagation in a cortical model. *Journal of Computational Neuroscience*, 22, Feb 2007.
- [24] K. Leslie, J. Sleight, M. Paech, L. Voss, C. Lim, and C. Sleight. Dreaming and electroencephalographic changes during anesthesia maintained with propofol or desflurane. *Anesthesiology*, 111(3):547–555, September 2009.

- [25] R. P. Lesser, S. H. Kim, L. Beyderman, D. L. Miglioretti, W. Webber, M. Bare, B. Cysyk, G. Krauss, and B. Gordon. Brief bursts of pulse stimulation terminate afterdischarges caused by cortical stimulation. *Neurology*, 53(9):2073–2081, December 1999.
- [26] D. T. Liley, P. J. Cadusch, and M. P. Dafilis. A spatially continuous mean field theory of electrocortical activity. *Network: Computation in Neural Systems*, 13:67–113, 2002.
- [27] D. T. Liley, P. J. Cadusch, and J. J. Wright. A continuum theory of electro-cortical activity. *Neurocomputing*, 26-27:795–800, 1999.
- [28] D. T. Liley and J. J. Wright. Intracortical connectivity of pyramidal and stellate cells: estimates of synaptic densities and coupling symmetry. *Network: Computation in Neural Systems*, 5:175–189, 1994.
- [29] B. A. Lopour and A. J. Szeri. *Advances in Cognitive Neurodynamics: Proceedings of the International Conference on Cognitive Neurodynamics*, chapter 86: Spatial Considerations of Feedback Control for the Suppression of Epileptic Seizures, pages 495–500. Springer, 2008.
- [30] B. A. Lopour and A. J. Szeri. A model of feedback control for the charge-balanced suppression of epileptic seizures. *Journal of Computational Neuroscience*, under review.
- [31] E. McKay, J. Sleigh, L. Voss, and J. Barnard. Episodic waveforms in the electroencephalogram during general anaesthesia: a study of patterns of response to noxious stimuli. *AIC*, accepted.
- [32] D. Minecan, A. Natarajan, M. Marzec, and B. Malow. Relationship of epileptic seizures to sleep stage and sleep depth. *SLEEP*, 25, 2002.
- [33] G. K. Motamedi, R. P. Lesser, D. L. Miglioretti, Y. Mizuno-Matsumoto, B. Gordon, W. R. S. Webber, D. C. Jackson, J. P. Sepkuty, and N. E. Crone. Optimizing parameters for terminating cortical afterdischarges with pulse stimulation. *Epilepsia*, 43(8):836–846, 2002.
- [34] M. Mourtazaev, B. Kemp, A. Zwinderman, and H. Kamphuisen. Age and gender affect different characteristics of slow waves in the sleep EEG. *Sleep*, 18(7):557–564, 1995.
- [35] T. I. Netoff, R. Clewley, S. Arno, T. Keck, and J. A. White. Epilepsy in small-world networks. *Neurobiology of Disease*, 24(37):8075–8083, September 2004.
- [36] E. Niedermeyer and F. L. da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2005.
- [37] R. Nieuwenhuys. The neocortex: an overview of its evolutionary development, structural organization and synaptology. *Anatomy and Embryology*, 190(4):307–337, Oct 1994.
- [38] P. L. Nunez and R. Srinivasan. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, 2006.

- [39] E. Olofsen, J. W. Sleight, and A. Dahan. Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. *British Journal of Anaesthesia*, 2008.
- [40] W. Penfield and H. Jasper. *Epilepsy and the functional anatomy of the human brain*. Little, Brown & Company, 1954.
- [41] M. Polito and P. Perona. Grouping and dimensionality reduction by locally linear embedding. *Neural Information Processing Systems Foundation*, 2001.
- [42] K. A. Richardson, S. J. Schiff, and B. J. Gluckman. Control of traveling waves in the mammalian cortex. *Physical Review Letters*, 94(028103), 2005.
- [43] L. S. Robblee and T. L. Rose. *Neural Prostheses: Fundamental Studies*, chapter 2: Electrochemical Guidelines for Selection of Protocols and Electrode Materials for Neural Stimulation. Prentice Hall, Englewood Cliffs, NJ, 1990.
- [44] R. Rocamora, J. C. Sánchez-Álvarez, and J. Salas-Puig. The relationship between sleep and epilepsy. *The Neurologist*, 14(65):S35–S43, 2008.
- [45] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- [46] S. T. Roweis and L. K. Saul. *Locally Linear Embedding*. World Wide Web, <http://www.cs.toronto.edu/~roweis/lle/>, accessed June 2009.
- [47] J. T. Rubinstein, F. A. Spelman, M. Soma, and M. F. Suesserman. Current density profiles of surface mounted and recessed electrodes for neural prostheses. *IEEE Transactions on Biomedical Engineering*, BME-34(11), November 1987.
- [48] L. K. Saul, S. T. Roweis, and Y. Singer. Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *Journal of Machine Learning Research*, 4:119–155, December 2003.
- [49] S. J. Schiff, X. Huang, and J.-Y. Wu. Dynamical evolution of spatiotemporal patterns in mammalian middle cortex. *Physical Review Letters*, 98(178102), 2007.
- [50] S. J. Schiff and T. Sauer. Kalman filter control of a model of spatiotemporal cortical dynamics. *Journal of Neural Engineering*, 5:1–8, 2008.
- [51] N. Spruston. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9(3):206–221, March 2008.
- [52] D. A. Steyn-Ross, M. L. Steyn-Ross, J. W. Sleight, M. T. Wilson, I. P. Gillies, and J. J. Wright. The sleep cycle modelled as a cortical phase transition. *Journal of Biological Physics*, 31:547–569, 2005.
- [53] M. L. Steyn-Ross, D. A. Steyn-Ross, and J. W. Sleight. Modelling general anaesthesia as a first-order phase transition in the cortex. *Progress in Biophysics & Molecular Biology*, 85:369385, 2004.

- [54] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, and D. T. J. Liley. Theoretical electroencephalogram stationary spectrum for a white-noise-driven cortex: Evidence for a general anesthetic-induced phase transition. *Physical Review E*, 60(6), 1999.
- [55] M. L. Steyn-Ross, D. A. Steyn-Ross, J. W. Sleigh, and D. R. Whiting. 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*, 68:021902, 2003.
- [56] M. F. Suesserman, F. A. Spelman, and J. T. Rubinstein. *In Vitro* measurement and characterization of current density profiles produced by nonrecessed, simple recessed, and radially varying recessed stimulating electrodes. *IEEE Transactions on Biomedical Engineering*, 38(5), May 1991.
- [57] F. T. Sun, M. J. Morrell, and J. Robert E. Wharen. Responsive cortical stimulation for the treatment of epilepsy. *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*, 5:68–74, January 2008.
- [58] S. Sunderam, N. Chernyy, J. Mason, N. Peixoto, S. L. Weinstein, S. J. Schiff, and B. J. Gluckman. Seizure modulation with applied electric fields in chronically implanted animals. *Proceedings of the 28th IEEE EMBS Annual International Conference*, 2006.
- [59] R. Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2), June 1972.
- [60] The Epilepsy Foundation. *About Epilepsy*. World Wide Web, <http://www.epilepsyfoundation.org/about>, accessed 2009.
- [61] M. T. Wilson, A. Steyn-Ross, J. W. Sleigh, M. L. Steyn-Ross, L. C. Wilcocks, and I. P. Gillies. The k-complex and slow oscillation in terms of a mean-field cortical model. *Journal of Computational Neuroscience*, 21:243–257, 2006.
- [62] M. T. Wilson, M. L. Steyn-Ross, D. A. Steyn-Ross, and J. W. Sleigh. Predictions and simulations of cortical dynamics during natural sleep using a continuum approach. *Physical Review E*, 72(051910), 2005.

Appendix A

Stochastic signals at relative step sizes

A.1 Generation of two-dimensional stochastic signals

To create the Brownian paths for a strip of cortex, we simply generate a matrix of Gaussian-distributed random numbers. The size of the matrix will depend on the desired number of space and time steps; for N time steps and M space steps, we can use the MATLAB command `ee = randn(N,M)`. In order to generate the same paths for each simulation, we reset the state of the random number generator using `randn('state',0)`. Note that each number in the `ee` matrix represents an *increment* of the Brownian path, as is needed for numerical integration with the Euler-Maruyama method (see equation 4.3 in [15]). To visualize the paths in time, we would need to look at the cumulative sum of the increments with a command like `plot(cumsum(ee))`. Another important distinction is that the paths represent Brownian motion in time, but each path is independent from its spatial neighbors (so each column of `ee` is an independent Brownian path, and each row contains independently generated random numbers). This means that we will need to treat space and time differently when scaling the signals.

Using equations (3.1) and (3.2) we can then generate two-dimensional stochastic inputs at relative step sizes. Here we verify these choices with a Matlab script:

```
% Generate matrix of random numbers
N = 60;
M = 40;
ee = randn(N,M);

% Plot 30th Brownian path vs. time; dt = 1
plot(1:N, cumsum(ee(:,30)));

% Generate path with dt = 2; plot again
ee2 = ee(1:2:end,:) + ee(2:2:end,:);
hold on
plot(1.5:2:N, cumsum(ee2(:,30)), ':'');

% Generate path with dt = 4; plot again
```

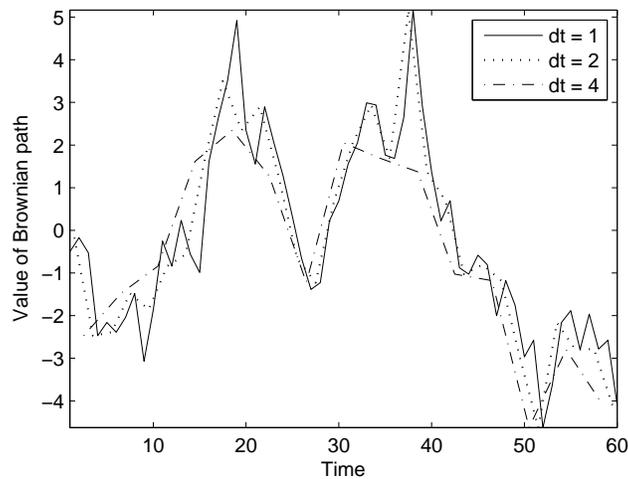


Figure A.1: Plot of the same Brownian path in time at three different step sizes.

```

ee3 = ee2(1:2:end,:) + ee2(2:2:end,:);
plot(2.5:4:N, cumsum(ee3(:,30)), '-.');
xlabel('Time')
ylabel('Value of Brownian path')
legend('dt = 1', 'dt = 2', 'dt = 4')

% Plot the stochastic input at 10th time step; dx = 1
figure
plot(1:M, ee3(10,:));

% Generate signal at dx = 2; plot again
ee4 = (ee3(:,1:2:end) + ee3(:,2:2:end))./sqrt(2);
hold on
plot(1.5:2:M, ee4(10,:), ':');

% Generate signal at dx = 4; plot again
ee5 = (ee4(:,1:2:end) + ee4(:,2:2:end))./sqrt(2);
plot(2.5:4:M, ee5(10,:), '-.');
xlabel('Space')
ylabel('Value')
legend('dx = 1', 'dx = 2', 'dx = 4')

```

This code generates equivalent stochastic inputs in space and time for three different step sizes. Figure A.1 shows the same Brownian path generated at three different values of Δt using this method, and Figure A.2 shows the stochastic signal over space at three different values of Δx .

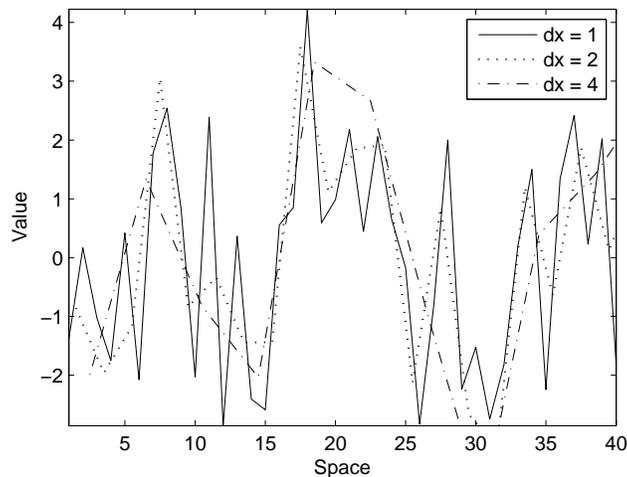


Figure A.2: Plot of the same stochastic input over a strip of cortex at three different step sizes.

A.2 Scaling preserves signal variance

In the spatial dimension, we will *not* scale the increments by $\sqrt{\Delta x}$ because each point in space is independent of all others. (Another way to think about this: if we did include this term the total scaling would be $\sqrt{\Delta t \Delta x}$, so reducing Δx would affect the variance of the signals in time even if Δt remained constant.) However, we will include a term for the *relative* scaling of the signal. We multiply by $1/\sqrt{2}$ every time we double the step size, as stated in equation (3.2); this ensures that the variance of the spatial signal is preserved as we increase Δx (relative to the *base* step size). We can verify this mathematically. Suppose we have M points in space, x_1, x_2, \dots, x_M with mean

$$m = \frac{1}{M} \sum_{i=1}^M x_i, \quad (\text{A.1})$$

and variance

$$m_2 = \frac{1}{M} \sum_{i=1}^M (x_i - m)^2. \quad (\text{A.2})$$

Now say that we generate the same signal over half as many points using the previously defined formula:

$$\tilde{x}_i \equiv \frac{1}{\sqrt{2}} (x_{2i-1} + x_{2i}), \quad i = 1, \dots, M/2. \quad (\text{A.3})$$

The mean of this new set of points is

$$\tilde{m} = \frac{1}{M/2} \sum_{i=1}^{M/2} \tilde{x}_i \quad (\text{A.4})$$

$$= \frac{2}{\sqrt{2}M} \sum_{i=1}^M x_i \quad (\text{A.5})$$

$$= \frac{2}{\sqrt{2}} m. \quad (\text{A.6})$$

And calculating the variance, we find

$$\tilde{m}_2 = \frac{1}{M/2} \sum_{i=1}^{M/2} (\tilde{x}_i - \tilde{m})^2 \quad (\text{A.7})$$

$$= \frac{2}{M} \sum_{i=1}^{M/2} \left(\frac{1}{\sqrt{2}} (x_{2i-1} + x_{2i}) - \frac{2}{\sqrt{2}} m \right)^2 \quad (\text{A.8})$$

$$= \frac{1}{M} \sum_{i=1}^{M/2} (x_{2i-1} - m + x_{2i} - m)^2 \quad (\text{A.9})$$

$$= \frac{1}{M} \sum_{i=1}^{M/2} (x_{2i-1} - m)^2 + (x_{2i} - m)^2 + (x_{2i-1} - m)(x_{2i} - m) \quad (\text{A.10})$$

$$= \frac{1}{M} \sum_{i=1}^M (x_i - m)^2 + \frac{1}{M} \sum_{i=1}^{M/2} (x_{2i-1} - m)(x_{2i} - m) \quad (\text{A.11})$$

$$= m_2 + \frac{1}{M} \sum_{i=1}^{M/2} (x_{2i-1} - m)(x_{2i} - m) \quad (\text{A.12})$$

Noting that $m = 0$ and that the independence of the x_i values causes the sum of the cross terms $x_{2i-1}x_{2i}$ to approach zero as $M \rightarrow \infty$, we obtain the desired result: $\tilde{m}_2 = m_2$. Therefore, the scaling term preserves the variance of the signal.

Another way to verify this is to use a large value of M in the MATLAB code from Appendix A.1 and compare the variances of the signals. With $M = 50,000$ we get $\text{var}(\text{ee3}(10, :)) = 3.9864$, $\text{var}(\text{ee4}(10, :)) = 3.9870$, $\text{var}(\text{ee5}(10, :)) = 4.0192$.

Appendix B

Stochastic signals at absolute step sizes

Although it is necessary to add the $1/\sqrt{2}$ scaling term in order to compare various values of Δx and verify strong convergence, this will not be necessary once adequate values of Δt and Δx have been chosen. In a typical simulation of the cortical SPDE's, the stochastic increments will only be scaled by $\sqrt{\Delta t}$; this means that the variance of the increments will be $\sigma^2 = \Delta t$. Although this variance changes with the step size, the *mean variance of the Brownian paths will remain constant*. This can be verified with a MATLAB script:

```
n = 50000;    % number of points
dt = .05;    % smallest time step
dx = .1;     % spatial step

% Generate 400 paths at different dt values and calculate the variance
for i = 1:400
    x1 = randn(1,n)*sqrt(dt); % generate random path with proper scaling
    x2 = randn(1,n/2)*sqrt(2*dt);
    x3 = randn(1,n/4)*sqrt(4*dt);
    x4 = randn(1,n/8)*sqrt(8*dt);
    var1(i) = var(cumsum(x1)); % find the variance of the Brownian paths
    var2(i) = var(cumsum(x2));
    var3(i) = var(cumsum(x3));
    var4(i) = var(cumsum(x4));
end

% Calculate the average variance for each step size
mvar1 = mean(var1);
mvar2 = mean(var2);
mvar3 = mean(var3);
mvar4 = mean(var4);
```

This script calculates 400 Brownian paths at 4 different step sizes, calculates the variance of each path, and then determines the mean variance for each step size. For example, one

execution of the script produced the results $\text{mvar1} = 37.5368$, $\text{mvar2} = 37.7847$, $\text{mvar3} = 38.4455$, $\text{mvar4} = 40.8610$. These numbers appear to converge as we increase the number of paths included in the mean.