UNIVERSITY OF CALIFORNIA, SAN DIEGO

Path Integral Techniques for Estimating Neural Network Connectivity

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

in

Physics

by

Christopher J. Knowlton

Committee in charge:

Professor Henry D.I. Abarbanel, Chair
Professor Timothy Gentner
Professor Michael Holst
Professor Julius Kuti
Professor Oleg Shpyrko

2014

The dissertation of Christopher J. Knowlton is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2014

TABLE OF CONTENTS

LIST OF SYMBOLS

Throughout this work an attempt will be made to use consistent notation for state vectors, measurements, and other common vector and matrix quantities.

$\mathbf{x}(t_i)$	- State vector at time t_i.
$\mathbf{x}$(i)	- State vector at time step i. This is equivalent to above, but for discrete time steps.
x_i	- i'th component of state vector $\mathbf{x}$.
	Time dependence is always assumed even if not explicitly stated.
$\mathbf{p}$	- Vector of time independent model parameters.
$\mathbf{f}(\mathbf{x},\mathbf{p})$	- Dynamical equation for the state vector $\mathbf{x}$. $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p})$
$\mathbf{F}(\mathbf{x}(t_i), \mathbf{p})$	- Descrete mapping of dynamics. $\mathbf{F}(\mathbf{x}(t_i), \mathbf{p}) = \mathbf{x}(t_i) + \int_{t_i}^{t_{i+1}} dt' \mathbf{f}(\mathbf{x}(t'), \mathbf{p})$
$\mathbf{h}(\mathbf{x})$	- observable of the system.
$\mathbf{y}$(t), $\mathbf{y}$(i), y_i	- measurement vector of observable quantity.
	Dimensions of y are typically much less than x.

$\mathbf{X}$	- Path vector. $\mathbf{X} = \{\mathbf{x}(0), \mathbf{x}(1), ..., \mathbf{x}(N), \mathbf{p}\}$
	Assumed to represent whole path all the time and all parameters
$\mathbf{Y}$(T)	- Prior measurements. $\mathbf{Y}(T) = \{\mathbf{y}(0), \mathbf{y}(1), ..., \mathbf{y}(T)\}$ where $T < N$

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor and committee chair Professor Henry Abarbanel. Without his support (and occasional prodding) this thesis would never have been completed.

I would like to acknowledge the Margoliash lab at University of Chicago for excellent neuron data and instruction on how that data is collected. Images from some of their experiments have been included with permission.

Many figures and tables in chapters 4, 5, 6, and 7 are reproductions of figures that appeared in C. Knowlton, C.D. Meliza, D. Margoliash, and H.D.I. Abarbanel, "Dynamical estimation of neuron and network properties III: network analysis using spike times" *Biological Cybernetics* **108**:3 261-273 (2014). The the thesis author is the primary author of this paper.

Figure 4.10 and part of Figure 6.3 are reproduced by permission of authors from Meliza, C. D., M. Kostuk, H. Huang, A. Nogaret, D. Margoliash,and H. D. I. Abarbanel, "Estimating parameters and predicting membrane voltages with conductance-based neuron models" *Biol. Cybernetics*, **108** 495-516 (2014)

I would like to acknowledge the Windandsea crew for providing sanity through Settlers, Dominion, and Smash.

VITA

B.A. Physics, Reed College 2008

Graduate Teaching Assistant 2009-2010, 2014

Graduate Research Assistant 2010-2014

Assistant Director of UCSD Young Physicists Program 2010-2011, 2013-2014

Director of UCSD Young Physicists Program 2011-2013

Ph.D. Physics, UCSD 2014

PUBLICATIONS

Christopher Knowlton, Henry D. I. Abarbanel, (2012) Estimating neuron and network properties using data assisted synchronization Poster, Dynamics Days 2012, Baltimore MD

Christopher Knowlton, Henry D. I. Abarbanel, (2012) Estimating neuron and network properties using data assisted synchronization Poster, SFN 2012, New Orleans LA

Christopher Knowlton and Henry D. I. Abarbanel, (2014) Estimating Network Architecture and Stimulus in Lotka Volterra Networks Poster, Dynamics Days 2014, Atlanta GA

C. Knowlton, C.D. Meliza, D. Margoliash, and H.D.I. Abarbanel, "Dynamical estimation of neuron and network properties III: network analysis using spike times" *Biological Cybernetics* **108**:3 261-273 (2014)

Christopher Knowlton "Estimating Network Properties from Spike Timing Information" Contributed Talk, MBI Workshop for Young Researches in Mathematical Biology, 2014, Columbus, OH

ABSTRACT OF THE DISSERTATION

Path Integral Techniques for Estimating Neural Network Connectivity

by

Christopher J. Knowlton

Doctor of Philosophy in Physics

University of California, San Diego, 2014

Professor Henry D.I. Abarbanel, Chair

Characterizing the behavior of networks of neurons requires accounting for the differing levels of measurements at different scales. At the single neuron level, intracellular recordings allow for highly accurate membrane potential measurements in response to an designed applied current. Because the probes used for the single neuron experiments are large compared to the cells themselves, these voltage measurements cannot be assumed to be available for any more than a few cells at a time. Instead of voltage measurements of the potential across the cell membrane, extracellular voltage measurements combined with spike sorting algorithms allow for measurements of spike times on orders of magnitude more neurons. This spike timing information provides much less information per neuron, requiring the development of new methods to estimate the states and connectivity of a network of neurons.

Previous work [Toth, et al., 2011, Kostuk, et al., 2012, Meliza, et al., 2014] has demonstrated the ability of a path integral formulation to characterize the behavior of individual neurons given time series voltage data. We expand on this to potential future experiments to characterize the behavior of synaptic connections, and other external currents acting on neurons and two possible means for determining the connectivity of a network of neurons given spike timing information.

Chapter 1

Introduction

Networks of neurons play a key role in the perception, interpretation, and response to external stimulus in complex organisms. Characterizing the behavior of a particular network of neurons requires identification of the number and type of the neurons in the cell, identification of the stimulus that acts on the network, and a determination of the synaptic connections between each cell. This is complicated by the sparsity of measurements of the underlying dynamics of the system.

This thesis will present a novel method to use an estimation procedure motivated by statistical physics to:

- Develop biophysically accurate models of individual neurons in isolation.
- Use isolated neuron models as a filter to develop models for the synaptic and other external currents acting on neurons in a network.
- Use individual neuron models and synaptic models as building blocks to create a framework that can be used to estimate network structure and behavior from sparse activity measurements.

We start with a development of a path integral formulation for estimating unmeasured states and parameters in non-linear dynamical system. We use a series of simple equations used in oceanographic and atmospheric modeling to demonstrate a relationship between dynamical control of a chaotic system and the ability to estimate unmeasured model states and parameters.

We demonstrate the use of this path integral formalism to estimate properties of isolated individual neurons through a multi step process that combines simulated experiments with controlled measurements of real neurons in vitro. Simulated experiments are used to determine the types of measurements and stimulus protocols that allow for resolution into unmeasured states and parameters for a given model. The effective stimulus protocols are then applied to a neuron in a slice from the high vocal center (HVC) and the resulting voltage response is used to fit a model that can be used to predict future behavior.

The neuron models developed through experiments on isolated neurons can be used as a passive filter to allow for readily available time series voltage measurements to be used to estimate the currents caused by the synaptic connections from nearby neurons. The estimated currents in conjunction with stimulated activity in nearby neurons is shown through simulated experiments to have the potential to allow for the properties of the various types of synaptic connections present in networks to be estimated.

Provided accurate models of individual neurons and synaptic activity the sparse information provided by extracellular recordings can potentially be translated into a series of stereotyped, disjointed, voltage and synaptic activity waveforms. We demonstrate that because these waveforms correspond to the regions of instability in the model, they are sufficient for estimating the full state of the neurons between measured windows. We show that the sparse measurements are sufficient to estimate the connectivity of small simulated networks of recurrently connected neurons.

If it can be demonstrated that network behavior and function is determined primarily based on activity level (firing rate) we demonstrate that the methods developed for spiking neurons can equally applied to rate networks. We use a model developed to describe olfaction to demonstrate that connectivity and applied stimulus can be estimated for these types of models given activity measurements of each node.

Chapter 2

Data Assimilation in Dynamical Systems

Before delving the behavior of neurons and networks we must first discuss the mathematical framework with which we will approach the problem. To this end we step back and look to what the general question we are trying to answer.

Given some time series of measurements of a system - what is the future behavior of that system?

We approach this problem by assuming the system is dynamical - that the system can be described by a rule that determines the change in the state as a function of the current state of the system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t)). \tag{2.1}$$

The assumption that the problem can be dealt with as a dynamical system changes the problem into two parts. First, given $\mathbf{f}$ and a set of measurements - determine the unmeasured states and time independent parameters at the end of the measured window. Then given that final state and estimated final state - integrate the model forward to predict future behavior.

For now we assume that both the functional form and parameters of $\mathbf{f}$ are known for each system of interest. Dealing with parameter estimation simultaneously

with state estimation is non-trivial, particularly when $\mathbf{f}(\mathbf{x})$ is non-linear. Estimating parameter values requires the ability to determine what $\mathbf{f}(\mathbf{x})$ *should* be at each point - which requires knowledge of the state at each point. We will therefore be postponing a discussion of parameter estimation until after developing the path integral method in the next chapter.

The measurements of the system - $\mathbf{y}(t)$ are either a subset of the dynamical variables -$\mathbf{x}(t)$ or functions of those variables - $\mathbf{h}(\mathbf{x}(t))$. Generally the number of measurable states will be less than the number of states, requiring some method to estimate the remaining states.

Implicit in this interpretation of the measurements is the assumption that the dynamics of the physical system evolves according to the rules of our model. This leads to a pair of dynamical systems evolving simultaneously - the physical system and the model. To accurately predict future behavior, the state of the model system at the end of the data set must be the same as the physical system. By the end of the available data the two systems must be synchronized.

The synchronization of a pair of dynamical systems depends on the structure of the model and the information provided by the measurements of the system. We first examine how the structure of the dynamics and the potential for chaotic behavior prevent estimation of unmeasured states. We then discuss how the use of regular measurements of a subset of states can counteract the estimation errors caused this behavior.

2.1 Chaos in Dynamical Systems

Depending on the dynamics of the system, perturbations in initial conditions will grow or shrink in size according to the current state of the system. We will treat the state and parameter estimations in a moment and first consider systems in which both the form and parameters of a model describing some system of interest are known. Absent measurements with fixed, known model parameters, a perturbation - $\delta\mathbf{x}$ will be stretched and compressed based on the dynamics of the system.

$$\delta\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x} + \delta\mathbf{x}, t) - \mathbf{f}(\mathbf{x}, t) \tag{2.2}$$

$$\begin{aligned} \delta\dot{x}_i &\approx \textstyle\sum_j \frac{\partial f_i(\mathbf{x},t)}{\partial x_j}\delta x_j \\ &= J_{ij}(\mathbf{x},t)\delta x_j \end{aligned} \tag{2.3}$$

Because these changes are time and state dependent, describing the long term evolution of this perturbation is dependent on the ergodicity of the system. For a ergodic continuous dynamical system this translates to orbits that are bounded and dense. Over sufficiently long periods of time this ergodicity causes time averages to become spatial averages weighted by an appropriate density function, allowing us to drop the time dependence from both the perturbation and Jacobian. [Oseledets, 1968] This suggests that for sufficiently long periods of time, the scaling of any minuscule perturbation will be independent of the initial state of the system provided the dynamics evolve to the correct basin of attraction.

$$\int_{t=0}^{T} dt J(\mathbf{x},t) \rightarrow \int_{\mathbf{x}} d\mathbf{x} J(\mathbf{x})\rho(\mathbf{x}) \tag{2.4}$$

where $\rho(\mathbf{x})$ is the probability of being in state $\mathbf{x}$ at some arbitrary time. On average:

$$\delta\dot{\mathbf{x}} = \left[\int_{\mathbf{x}} d\mathbf{x} J(\mathbf{x})\rho(\mathbf{x})\right]\delta\mathbf{x}, \tag{2.5}$$

which is an exponential function in time dominated by the largest eigenvalue of the average Jacobian.

$$\lambda_{max} = \lim_{T\rightarrow\infty}\lim_{\delta\mathbf{x}(0)\rightarrow 0} \ln\frac{|\delta\mathbf{x}(T)|}{|\delta\mathbf{x}(0)|} \tag{2.6}$$

In practical terms, this growth in a perturbation corresponds to the creation of information about the initial state of the system. As closely spaced initial conditions diverge, comparing the evolution of two trajectories over time gives an increasing amount of information about the initial state of the system. It is for this reason that chaotic systems are said to 'create' information, while 'regular' ($\lambda_{max} < 0$) systems 'destroy' information. For predicting future behavior, one would need measurements of the system to discover the additional information created by chaos whereas in a regular system no measurements are needed to find the current state of the system because all of the information about the initial conditions has been lost. In practice we are typically interested in using a model to predict future behavior and in such cases insensitivity to initial conditions is generally a positive feature.

For a continuous dynamical system, if the perturbation is in the direction of the evolution of the system (forwards or backwards) the perturbation amounts to time delay in the dynamics of the system. Since this perturbed path must follow the same trajectory (albeit passing the same points in state space at different times) the size of the perturbation averaged over long periods of time must be constant. Therefore all continuous dynamical system must have a null vector corresponding to motion along the normal dynamics. When we refer to the largest global Lyapunov exponent, it is always a reference to the largest non-zero exponent. The existence of a null vector has obviously no bearing on any sort of synchronization or divergence argument - but consequently there cannot be any continuous chaotic systems of dimension less than 3.[Poincaré, 1892] A chaotic systems by definition approaches arbitrarily close to every point in a region of state space in a finite time. This region forces the dynamics to be bounded, and by extension globally dissipative. The trace of the Jacobian in a dissipative system must be negative and the sum of eigenvalues is equal to the trace. When combined with the presence of a null eigenvalue, for a system to have a positive Lyapunov exponent it must also have a larger magnitude negative Lyapunov exponent - hence a third state is needed.

2.2 Calculating Lyapunov Exponents

Recall that for ergodic systems, the local Lyapunov exponents relate to the global exponents by way of a geometric average expressed through a mapping derived from the dynamics.

$$\begin{aligned} \mathbf{x}(t+dt) &\approx \mathbf{x}(t) + dt\mathbf{f}(\mathbf{x}(t)) \\ \partial x_i(t+dt)/\partial x_j(t) &= \delta_{ij} + dtJ_{ij}(\mathbf{x}(t)) \\ \delta x_i(t+dt) &= \delta x_j(t)\big(\delta_{ij} + dtJ_{ij}(\mathbf{x}(t))\big) \end{aligned} \tag{2.7}$$

This allows the final perturbation to be expressed in terms of the initial perturbation:

$$\delta x_i(T) = \delta x_j(0) \prod_{k=0}^{N} [\delta_{ij} + dtJ_{ij}(\mathbf{x}(kdt))] \tag{2.8}$$

where T = N dt. For large dimensional systems this calculation is done through directly perturbing the initial conditions for a range of perturbations to find the

largest value. However, for small[1] dimensional systems we can calculate all of the the Lyapunov exponents directly.

If we treat the mapping $A_{ab}(k)$ to be the evolution of perturbation from one step to the next:

$$\delta x_a(t_{k+1}) = A_{ab}(t_k)\delta x_b(t_k) \tag{2.9}$$

$$A_{ab}(k) = \delta_{ab} + dtJ_{ab}(\mathbf{x}(t_k)), \tag{2.10}$$

we can decompose A through QR decomposition such that:

$$\mathbf{A}(k)\mathbf{Q}(k-1) = \mathbf{Q}(k)\mathbf{R}(k). \tag{2.11}$$

where $\mathbf{Q}(0) = \mathbf{I}$. Thus:

$$\mathbf{A}(N)\mathbf{A}(N-1)...\mathbf{A}(1) = \mathbf{Q}(N)\mathbf{R}(N)\mathbf{R}(N-1)...\mathbf{R}(1). \tag{2.12}$$

Since $\mathbf{R}(k)$ are all right triangular matricies the eigenvalues can be easily calculated by finding the time average of the log of the diagonal elements of R: [Abarbanel, 1996]

$$\lambda_a = \lim_{T\to\infty} \frac{1}{Tdt} \sum_{k=1}^{T} \log[R_{aa}(k)] \tag{2.13}$$

which yields reasonable approximations for finite values of T. As N goes to infinity, $\mathbf{Q}$[N] goes to $\mathbf{1}$ and thus will not contribute to the sum of logs in the limit. Note the dt and T do not cancel as T goes to infinite time irrespective of the step size. The step size changes the rate at which the matrix R changes such that the geometric average of R in some interval remains constant. The factor of dt accounts for this averaging.

2.3 Regular Attractors - FitzHugh-Nagumo

While we have demonstrated that small perturbations in certain systems will either converge or diverge, generally our model will not start near the same point as the physical system. Because the physical systems we are interested in tend to be dissipative, we can expect that the orbits of the system over time to converge into

[1]'Small and large' are of course a function of computing power. Here small is taken to mean a calculation that can be made on a desktop computer in order minutes.

some bounded region. The size and location of this bounded region - the attractor - will provide us with the limits with which can predict future behavior of the system.

The most simplistic attractor is the fixed point. Many dynamical systems have a state for which $\mathbf{f}= 0$ and if this point occurs in a region where the dynamics are regular, it will also be an attractor. The convergence of nearby trajectories - one of which is at rest is equivalent to saying that the point is stable with respect to perturbations. Note that not all fixed points are attractors - the inverted pendulum has a fixed point when the pendulum is vertical, but is extremely sensitive to perturbations.

While all trajectories in the basin of attraction converge to a fixed point, there is no reason that they will always converge to each other at all times. Since regularity and chaos are global, not local, properties of an orbit, the system may make large diversions and exhibit *local* unstable behavior even if both trajectories are in the same basin of attraction. Figure 2.1 shows a plot of several trajectories of the FitzHugh-Nagumo neuron model (FHN) [FitzHugh, 1955]:

$$\dot{x_1} = x_1 - \frac{x_1^3}{3} - x_2 + I_{ext} \tag{2.14}$$

$$\tau\dot{x_2} = x_1 + a - bx_1 \tag{2.15}$$

where a = 0.7, b = 0.8, τ = 12.5, and I_{ext} is the driving stimulus - here chosen to be constant in time. While all trajectories converge to the same resting state (fixed point), some trajectories include large excursions.

The FHN model is a qualitative approximation of more complex, biologically motivated models such as Hodgkin-Huxley type models that we will be using in this thesis (see Chap. 4). Generally the state of x_1 is equivalent to the voltage across the cell membrane and x_2 to the excitability of that membrane, however the values do not directly map to any physical quantity. The dynamics of the FHN model are designed to capture the spiking behavior of neurons and other excitable systems by producing characteristic excursions in state space in response to sufficiently strong positive applied stimulus. For now we are interested more in its dynamical structure than the system it describes

As the stimulus I_{ext} increases, the resting state approaches the threshold that separates the two types of trajectories. (black and red vs green and blue in Figure 2.1) At I_{ext} =?, the attractor morphs from a fixed point into a stable periodic orbit - called

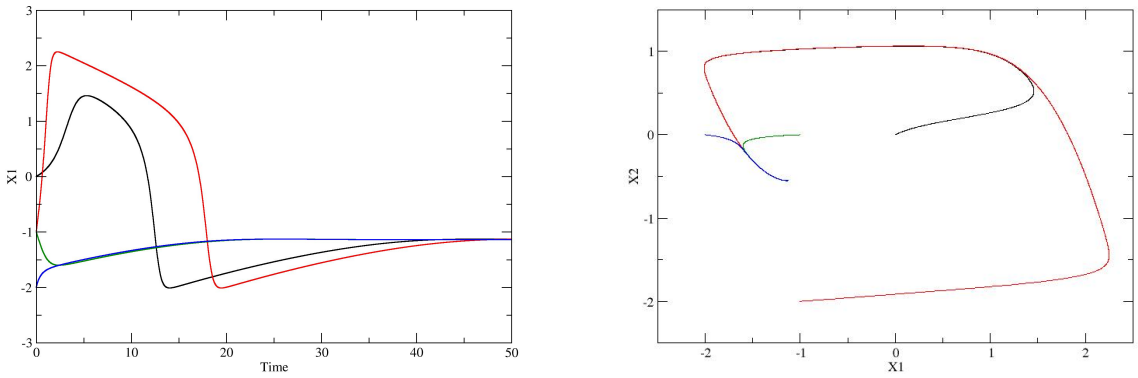

Figure 2.1: Plot of x_1 vs time (left) and phase plot (x_2 vs x_1, right) for the FitzHugh Nagumo model with constant ($I_{ext} = 0.1$) driving stimulus. While all trajectories converge to the same fixed point at (-1.15, 0.55), local instabilities result in temporary divergent behavior.

a limit cycle or relaxation oscillation. This transition, a Hopf bifurcation, will play in an important part of our ability to estimate properties of this and other more complex neuron models later in the thesis (see 6.1).

Figure 2.2 shows the same FHN model with a constant stimulus of 0.5. All trajectories converge to the same limit cycle attractor but do not converge to the same phase on that attractor. Additional information from measurements will be needed to estimate that phase.

While evolution to a fixed point or limit cycle provides us with increasing information about the state of the system, it appears it would be faster to simply guess that the system is at the location of that fixed point or on its limit cycle. The key of this convergence comes into play when the system is evolving between a series of fixed points or limit cycles as the dynamics changes - such as when a time varying forcing acts on the system.

As was demonstrated in Figs. 2.1 arbitrary initial conditions converge to the same fixed point or limit cycle. Given that all trajectories are (eventually) converging to this fixed point they must also (eventually) converge to each other. Even if the location of the fixed point changes over time, the trajectories may still converge towards each other even if the current fixed point is never reached. Fig. 2.3 shows the

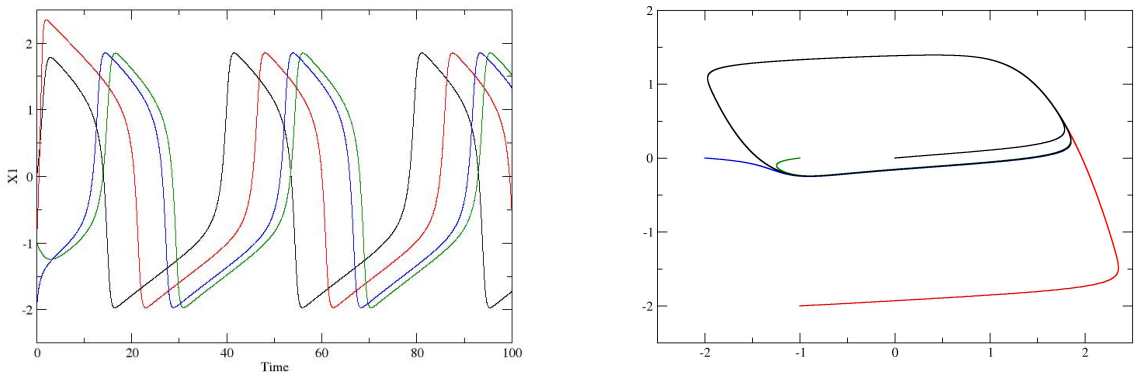

Figure 2.2: Plot of x_1 vs time (left) and phase plot (x_2 vs x_1, right) for the FitzHugh Nagumo model with constant ($I_{ext} = 0.5$) driving stimulus. All trajectories converge to the same limit cycle, but are out of phase.

evolution of a driven FHN system with a known stimulus. Even though the system passes between regions exhibit limit cycle behavior and fixed point behavior the orbits will converge because the dynamics are always regular.

It is possible to contrive a stimuli that results in chaotic behavior in the FHN system. Recall that chaos is ultimately the average effect over the local (in)stability that the system encounters in its orbit. The FHN model has a region of instability corresponding to the Hopf bifurcation point separating the limit cycle and fixed point behavior. By providing periodic forcing at around the same frequency of that limit cycle the system spends a larger fraction of its orbit in this unstable region [Chou, 1996]. Of course adding this periodic forcing (or other time varying forcing) is essentially adding additional dynamical states to the system. Thus even thought the system is nominally 2D and should not be chaotic per Poincare - adding explicit time dependence can lead to chaotic orbits. The local instabilities caused by a Hopf bifurcation will be revisited when we discuss estimating properties of the Hodgkin-Huxley neuron model.

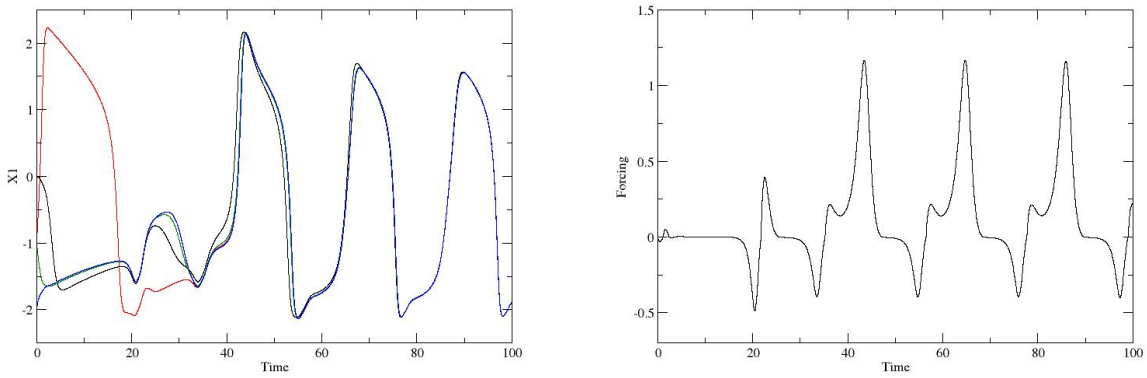

Figure 2.3: Plot of x_1 vs time (left) for several orbits on a FHN oscillator with known forcing (right). Despite the location and type of the attractor changing over time all trajectories converge.

2.4 Strange Attractors - Lorenz 63

For systems with chaotic behavior, closely spaced initial conditions will diverge over time while the dynamics are overall dissipative. This results in orbits converging to some region of state space with all orbits in that region (or any sub region) smearing out over the full region over time. To examine the properties of this 'strange' attractor we turn to the Lorenz 63 oscillator [Lorenz, 1963].

$$\begin{aligned} \dot{x}_1 &= \sigma(x_2 - x_1) \\ \dot{x}_2 &= -x_2 + rx_1 - x_1x_3 \\ \dot{x}_3 &= -bx_3 + x_1x_2 \end{aligned} \tag{2.16}$$

where $\sigma = 16$, $r = 40$, and $b = 1$.

As can be seen in Figure 2.4 (left) closely spaced initial conditions on the attractor of the Lorenz 63 system diverge over time while simultaneously being confined to a finite region of state space. Using the methods in Appendix 2.2, we find that the Lyapunov spectrum for the Lorenz system are {0.52, 0, -18.6}. Note that as should be expected for a dynamical system, the trace of the Jacobian is equal to the sum of the Lyapunov exponents and there is a null value corresponding to perturbations in

Figure 2.4: Left: plot of x_2 vs x_1 for three arbitrary initial states outside the attractor. Right: plot of x_3 vs x_1 for a pair of closely spaced initial conditions (grey point - near center) on the attractor. The global dissipative properties of the dynamics insure that orbits starting outside the attractor converge towards the attractor - but the presence of a positive Lyapunov causes small regions on the attractor to diverge to that same attractor.

the direction of motion.

Despite initial conditions diverging, this system still has an attractor due to the dissipation of the system. Figure 2.4 (right) shows that orbits starting outside the region in the left hand image will all converge into the same region state space. However, unlike the fixed point and limit cycle where the dynamics are confined to a 0 or 1 dimensional surface, respectively, the orbit on the chaotic attractor is smeared across a fractal surface in state space. Thus, even though the dynamics become confined to some smaller dimensional surface, the divergence of closely spaced orbits destroys any information about where we are on that surface over a time scale of $1/\lambda_{max}$. To determine where the state of a chaotic system is, our model system must be provided with additional information about its current location on that attractor.

2.5 Coupled Systems - Data Assisted Synchronization

Predicting future behavior of some physical system of interest requires the knowledge of the full state of the system. This state can potentially be found by evolving an identical, model system in parallel with the physical system until they synchronize. For non-linear systems, chaotic behavior and forced switching between different attractors create complicates our ability to synchronize. Given the possibility to periodically measure components of the system, can we use this information to assist in synchronizing our model with the physical system?

2.5.1 The Twin Experiment

Since we wish to the two trajectories to converge we can attempt to couple the two systems together by introducing a forcing term to the model.

$$\dot{x_i} = f_i(\mathbf{x}, \mathbf{p}) + K_{ij}(y_j - h_j(\mathbf{x})). \tag{2.17}$$

$$y_j = h_j(\mathbf{x}') + \eta_j(\mathbf{x}', t) \tag{2.18}$$

Where the measurements - $\mathbf{y}$, are a known function $\mathbf{h}(\mathbf{x})$ of the state of the physical system $\mathbf{x}$' with additive (or possibly multiplicative) noise η due to technical characteristics of the measurement tools. K_{ij} will generally be proportional to $\partial h_i(\mathbf{x})/\partial x_j$ as it maps the effect of a change in state j into the expected measurement i.

In principle $\mathbf{h}(\mathbf{x})$ can be an arbitrary function of the dynamical variables. However, for we will limit ourselves to systems whose observables are individual states - $h_i(\mathbf{x}) = x_i$. In this case K_{ij} is a positive diagonal matrix with L non-zero elements where L is the number of measured states ($L < D$). For this reason we will now refer to this coupling with a single index $K_i = K_{ij}$ if $i = j$, 0 else.

We might naively think that given measurements of some subset of the dynamical variables, we can simply replace the measurable variables with the measurements, equivalent[2] to $K_i \to \infty$. There are several reasons why this should not be done. First, the dynamics of the system are continuous while the measurements are discrete in

[2]modulo an implicit time delay and a very stiff equation to integrate

time. Unless the measurement frequency is similar to the integration step this leads to a need to interpolate between measurements - essentially requiring the full state of the system. Second, the measurements of the system will be wrong. Noise in the measurement device will lead to us forcing the system to incorrect states. The maximum size of the coupling is limited by the accuracy of the measurements. (See 3.2)

Third and most important, our model is wrong. While we have been assuming a 'perfect' model, the reality is that the small dimensional neuron models we are interested in are approximations of the statistical properties of billions of individual molecules in a heat bath - the fact that there exists some small dimensional model that even roughly captures some component of the behavior of the system is a minor miracle. Given that we are introducing a coupling to what is essentially a different system, this coupling will be smoothing over the inaccuracies in the model. For this reason we want the size of the coupling to measurements to be as large as needed to synchronize the system but no larger.

To determine how large a coupling, and to what, is appropriate we use two identical model systems. This eliminates the possibility the models are wrong and allows us to determine appropriate measurement routines under ideal circumstances. While this does not suggest that a particular protocol will work on a real system, if coupling to certain states does not work under idealized conditions it will certainly not work when there are uncertainties in the model.

2.5.2 Conditional Lyapunov Exponents

The introduction of a coupling term leads to a modification of the effective Jacobian of the system:

$$J'_{ij}(\mathbf{x}) = J_{ij}(\mathbf{x}, t) - K_i \tag{2.19}$$

where we are assume that $h_i(\mathbf{x}) = x_i$ for measured states. Note that in this case, coupling will appear on the diagonal of the modified Jacobian. The Lyapunov spectrum of this coupled spectrum will thus be more negative than before, but the addition of an effectively time dependent stimulus will mean that there is no longer a null Lyapunov exponent. Because the Lyapunov spectrum of the system is being modified by

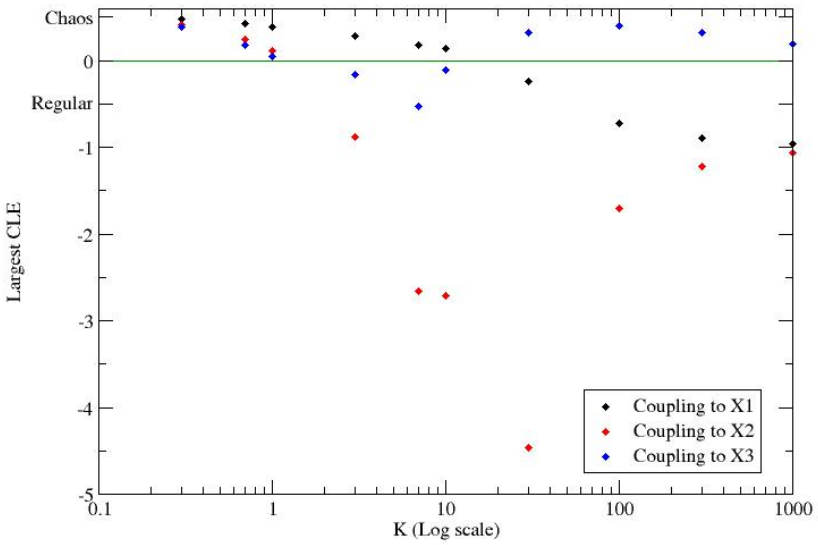

Figure 2.5: Plot of least negative conditional Lyapunov exponent (CLE) for Lorenz 63 system for coupling to each of the three variables. coupling to x_1 and x_2 result in synchronized dynamics for sufficiently strongly coupled systems, but coupling to x_3 will *usually* not result in synchronization.

the presence of this coupling they are referred to as conditional Lyapunov exponents (CLEs).

We use the procedure in 2.2 to calculate the CLE for the Lorenz 63 model (Eq. 2.17) with various coupling strengths to each of the three state variables. Coupling in measurements to either the x_1 or x_2 states will result in regular behavior provided sufficiently large coupling strength. However, coupling to the x_3 state will not cause the two systems to synchronize except for couplings of $\sim 5 - 15$. The existence of this region is potentially a numerical artifact - but it presses the point that the choice of what to measure and how strongly to couple those measurements is non-trivial.

For a small, three state system that one measured quantity is sufficient to estimate the full state of the system. However as the size and complexity of the system increases it is inevitable that a single measurement is insufficient to synchronize the system. For this question we examine the Lorenz 96 system, an atmospheric model on

cyclical set of states. This model is easily scalable in dimension with simple dynamics for each variable.

$$\dot{x}_i = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F \tag{2.20}$$

The index, i, ranges from 1 to N where $N \geq 5$, and is cyclical in N: $N + 1 \rightarrow 1, 0 \rightarrow N$. F represents a constant forcing term with chaotic behavior when F is greater than about 8 for all dimensions. We will work with an implementation of this model with F = 8.17, is found to have Lyapunov spectrum of about $\lambda = \{0.46, 0, -0.37, -1.3, -3.7\}$

Coupling to a measurement of just the x_1 variable is insufficient to lead to synchronization in the overall system. Because the Lorenz 96 model is symmetrical with respect to cycling indexes, any choice of measurement of a single variable is equivalent to any other. That a single measurement does not result in synchronization is a clear indication that additional information is needed from the measurements to synchronize the system. Adding in a second coupling to a measurement of x_4 provides sufficient information to synchronize the systems for N = 5. The CLE as a function of coupling strength are plotted in figure 2.6.

While coupling to certain states will result in synchronization over long periods of time, we are often interested in determining the states and parameters of a system in some finite window. The next chapter will present a method for using the information provided by this synchronization as a tool for determining the most likely path through state space for a set of measurements.

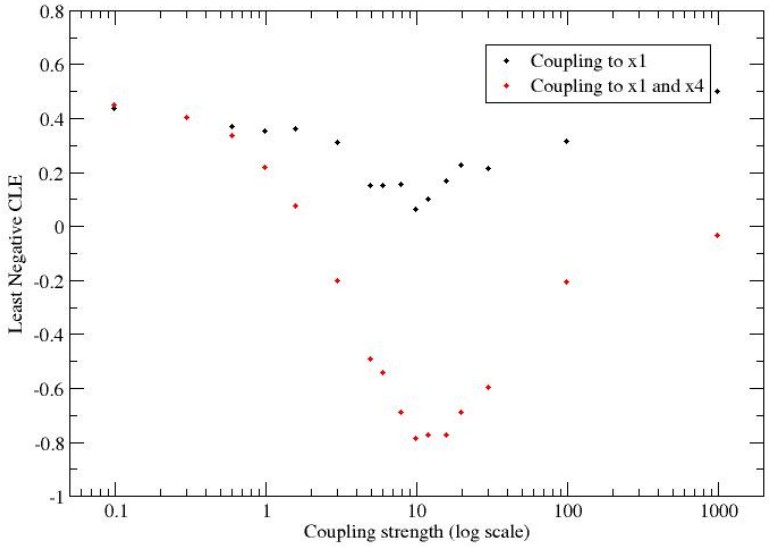

Figure 2.6: Least negative conditional Lyapunov exponent vs coupling strength in log scale for the D=5 Lorenz 96 system. For coupling to a single variable, insufficient information is passed between the coupled systems to result in synchronization regardless of the strength of the coupling. Coupling to two variables results in synchronization for sufficiently strong levels of coupling. The apparent convergence to $\lambda = 0$ may be an artifact of the inherent stiffness of the calculation of the CLE.

Chapter 3

Path Integral Method

While estimates of the final state would allow for prediction of future states in a known model, generally we will not start with an accurate model. This chapter will summarize the development of a path integral method that leverages the synchronization provided by appropriate regular measurements to estimate unmeasured model states and parameters.

3.1 Formulation of an Action

In addressing the general problem of transferring information from measurements we begin with L-dimensional observations $y_l(t_n)$; $l = 1, 2, ...L$; made at times t_n in an observation window $[t_0, t_1, t_2, ..., t_n, ...t_m = T]$. We want to communicate the information in these observations to a physical model in D-dimensional space governed by $\mathbf{x}(t_{n+1}) = \mathbf{F}(\mathbf{x}(t_n))$. Usually $D > L$ so we must estimate values of the unobserved state variables as well as values for any fixed parameters in the model. We have established [Abarbanel, 2013] an exact representation for the probability distribution for the state of the model at time t_m $\mathbf{x}(t_m) = \mathbf{x}(m)$ conditioned on measurements up until t_m; $\mathbf{Y}(m) = \{\mathbf{y}(0), \mathbf{y}(1), ..., \mathbf{y}(m)\}$. We denote this conditional probability distribution as $P(\mathbf{x}(m)|\mathbf{Y}(m))$.

Using this distribution function to describe noisy measurements and models with errors, we have given a precise formula for the expected value of any function $G(\mathbf{X})$ along the path through the observation window $\mathbf{X} = \{\mathbf{x}(0), \mathbf{x}(1), ..., \mathbf{x}(m)\}$ in

model state space conditioned on the measurements $\mathbf{Y}(m)$:

$$E[G(\mathbf{X})|\mathbf{Y}] = \frac{\int d\mathbf{X} \exp[-A_0(\mathbf{X},\mathbf{Y})]\, G(\mathbf{X})}{\int d\mathbf{X} \exp[-A_0(\mathbf{X},\mathbf{Y})]} \tag{3.1}$$

The action $A_0(\mathbf{X},\mathbf{Y})$ is composed of terms involving the observations, terms associated with the propagation of the model state from $\mathbf{x}(n)$ to $\mathbf{x}(n+1)$ and a term representing the initial distribution of the model state $P(\mathbf{x}(t_0)) = P(\mathbf{x}(0))$:

$$\begin{aligned} A_0(\mathbf{X},\mathbf{Y}) &= -\sum_{n=0}^{m} CMI(\mathbf{y}(n),\mathbf{x}(n)|\mathbf{Y}(n-1)) \\ &- \sum_{n=0}^{m-1} \log\Big[P(\mathbf{x}(n+1)|x(n))\Big] - \log[P(\mathbf{x}(0)] \end{aligned} \tag{3.2}$$

The measurement term involves the conditional mutual information of the L-dimensional measurement $\mathbf{y}(n)$ and the D-dimensional model state $\mathbf{x}(n)$ conditioned on all earlier measurements $\mathbf{Y}(n-1)$ [Fano, 1961]:

$$\begin{aligned} &CMI(\mathbf{y}(n),\mathbf{x}(n)|\mathbf{Y}(n-1)) = \\ &\log\left[\frac{P(\mathbf{y}(n),\mathbf{x}(n)|\mathbf{Y}(n-1))}{P(\mathbf{y}(n)|\mathbf{Y}(n-1))\, P(\mathbf{x}(n)|\mathbf{Y}(n-1))}\right] \end{aligned} \tag{3.3}$$

The dynamics taking the model state at t_n $\mathbf{x}(n)$ into the model state at t_{n+1} $\mathbf{x}(n+1)$ is represented in the transition probability $P(\mathbf{x}(n+1)|\mathbf{x}(n))$.

In the ratio defining the expected value of a function $G(\mathbf{X})$ over the path, $\mathbf{X}$ terms independent of $\mathbf{X}$ cancel. This leaves the action effective in estimating the expected values of functions on the state space path as

$$\begin{aligned} A_0(\mathbf{X},\mathbf{Y}) &= -\sum_{n=0}^{m} \log\Big\{P(\mathbf{y}(n)|\mathbf{x}(n),\mathbf{Y}(n-1))\Big\} \\ &- \sum_{n=0}^{m-1} \log\{P(\mathbf{x}(n+1)|\mathbf{x}(n))\} - \log P(\mathbf{x}(0)) \end{aligned} \tag{3.4}$$

which is effective in evaluating any expected value along the path in state space. A full derivation and discussion of this path integral can be found in [Abarbanel, 2013, Quinn & Abarbanel, 2010, Toth, et al., 2011].

Two general approaches to the evaluation of this integral have been explored. The first, which we pursue here, seeks a stationary path estimation, expanding about

minima of the action $A_0(\mathbf{X})$ [Laplace, 1774] where:

$$\frac{\partial A_0(\mathbf{X})}{\partial \mathbf{X}} = 0 \tag{3.5}$$

and the second is to use a Monte Carlo method [Metropolis, et al., 1953, Hastings, 1970, Quinn & Abarbanel, 2010] to sample the probability distribution $\exp[-A_0(\mathbf{X})]$. We do not write the argument $\mathbf{Y}$ from now on.

The stationary path approach is a numerical optimization problem. We use publicly available software packages for this purpose, SNOPT [Gill, et al., 2005] and IPOPT [Wächter & Biegler, 2006]. The solution to the variational problem yields an optimal path but does not give the fluctuations about that path. Those variations can be evaluated through corrections to the stationary path or through the direct use of a Monte Carlo evaluation of the integral, Eq. (1).

To use the variational principle, we must specify how the terms in the action $A_0(\mathbf{X})$ are to be approximated. This requires a choice about the errors in the measurements, a selection of how one represents model errors, and a statement about one's knowledge of the state $\mathbf{x}(0)$ when measurements begin.

In the first element of $A_0(\mathbf{X})$ we make the common assumption that the errors in measurements at time t_n are independent of measurements at earlier times $\mathbf{Y}(n-1)$ and are Gaussian. This is not a necessary assumption, but it suffices in the examples we address. We then write:

$$\begin{aligned} &-\log\Big\{P(\mathbf{y}(n)|x(n),\mathbf{Y}(n-1))\Big\} \\ &\approx \sum_{l=1}^{L}\Big\{\frac{R_m(l,n)}{2}(x_l(n)-y_l(n))^2\Big\} \end{aligned} \tag{3.6}$$

This also assumes the observation function $h_l(\mathbf{x}(n)) = x_l(n)$. The quantity $R_m(l,n)$ represents the RMS measurement error as $\sqrt{R_m^{-1}}$. The time dependence in $R_m(l,n)$ is introduced to allow us to turn this term on when observations are made and to turn it off when observations are absent.

The second term comes from the assumption that the physical model of the system is Markov [Abarbanel, 2013], meaning that the state $\mathbf{x}(t_{n+1}) = \mathbf{x}(n+1)$ depends only on the state at the previous time. We express the dynamics in the

deterministic, no model error, case as $\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{p})$. $\mathbf{p}$ is a set of fixed parameters in the model. In this setting the transition probability density is

$$P(\mathbf{x}(n+1)|\mathbf{x}(n)) = \delta^D(\mathbf{x}(n+1) - \mathbf{f}(\mathbf{x}(n), \mathbf{p})) \tag{3.7}$$

When we have model errors, we must reduce the state space resolution by broadening the delta function. There are many ways to do that, and here we replace the delta function by a Gaussian of a width of order $\sqrt{R_f^{-1}}$, so up to constants, we write

$$\begin{aligned} &-\log[P(\mathbf{x}(n+1)|\mathbf{x}(n))] = \\ &\sum_{a=1}^{D}\left\{\frac{R_f(a)}{2}(x_a(n+1) - x_a(n) - F_a(\mathbf{x}(n), \mathbf{p}))^2\right\} \end{aligned} \tag{3.8}$$

Formally, in the limit that $R_f \to \infty$, we recover the delta function of the deterministic case which has no model errors.

Without any prior knowledge of the distribution of states of the system when measurements commence, the probability distribution of initial conditions is assumed to be uniform over the range of states, so $-\log[P(\mathbf{x}(0))]$ is a constant that cancels in all expected values. The approximate action for the path integral is now expressed as:

$$\begin{aligned} A_0(\mathbf{X}) = &\sum_{n=0}^{m}\sum_{l=1}^{L}\left\{\frac{R_m(l,n)}{2}(x_l(n) - y_l(n))^2\right\} \\ &+ \sum_{n=0}^{m-1}\sum_{a=1}^{D}\left\{\frac{R_f(a)}{2}(x_a(n+1) - x_a(n) - F_a(\mathbf{x}(n), \mathbf{p}))^2\right\} \end{aligned} \tag{3.9}$$

Once the action is given, distributions on states and parameters can be found using Monte Carlo methods and the most probable distribution can be found using the variational method. This paper will focus on using the variational method to find minima of the action using the nonlinear optimization software IPOPT [Wächter & Biegler, 2006, Toth, et al., 2011, Kostuk, et al., 2012]. See appendix A for technical information on implementation of this method.

3.2 The Variational Step and Synchronization

Even with our Gaussian approximation to the action, the integrals needed to calculate expected values of states and parameters are intractable. Instead we rely on computational tools such as the variational method and/or Monte Carlo to sample the distribution in the regions where the path contributes non-negligibly to the integral. Since these regions are where the action is smallest - finding the minima of A_0 is equivalent to finding states and parameters near the 'correct' values.

Generally the action Eq. 3.9, can be minimized using a gradient descent, such as through the Langevin equation [Langevin, 1908]:

$$\frac{d\mathbf{X}}{ds} = -\nabla\partial A_0(\mathbf{X}) + \sqrt{2}\eta(0,1) \tag{3.10}$$

where s is a continuous perturbation variable and eta is a Gaussian process with mean 0 and variance α. This corresponds to Fokker-Planck equation [Fokker, 1914, Kolmogorov A., 1931] that describes the evolution of some distribution of initial paths.

$$\frac{d}{ds}P(\mathbf{X},s) = -\nabla_{\mathbf{X}}\left(P(\mathbf{X},s)\frac{\partial A_0(\mathbf{X})}{\partial \mathbf{X}}\right) + \nabla^2_{\mathbf{X}}P(\mathbf{X},s) \tag{3.11}$$

This equation has a steady state solution as $s \to \infty$ given by $P(\mathbf{X}) = \exp(-A_0(\mathbf{X}))$. Because of this property, evolving any initial guess of a path within the basin of attraction of the corresponding Langevin equation leads to the distribution given by $\exp[-A_0(\mathbf{X})]$.

If we drop the stochastic component of the Langevin equation (Eq. 3.10) the evolution of the path becomes a deterministic gradient descent. This gradient descent is a set of ordinary differential equations. Because A_0 contains $\mathbf{f}(\mathbf{x})$, which is non-linear, the gradient of A_0 will also contain non-linear terms. This suggests that we should be checking what the attractors of these dynamics look like. If the resulting dynamics have a single fixed point over all initial conditions, this fixed point is the global minimum of the action. If the gradient descent is chaotic over some region, we can expect to find a fractal number of local minima - thus making the variational method unable to estimate the true unmeasured states and parameters.

The actual dynamics of this gradient can be expanded out to first order in

dynamical time, t (as opposed to variational 'time' - s):

$$\frac{dx_i(t)}{ds} \approx \sum_{j=1}^{D} \left[\delta_{ij}+dtJ_{ij}(\mathbf{x}(t))\right] Rf_j(t) \left[x_j(t+dt)-x_j(t)-dtf_j(\mathbf{x}(t))\right]+Rm_i(y_i(t)-x_i(t)). \tag{3.12}$$

We immediately note that down to scaling factor, this is similar to the growth in a perturbation at each time step:

$$\Delta\mathbf{x}_i(t+1) = \sum_{j=1}^{D} \Delta x_j(t) \left[\delta_{ij} + dtJ_{ij}(\mathbf{x}(t))\right] - dtK_i\Delta y_i(t) \tag{3.13}$$

which is a combination of equations 2.8 and 2.19 where $\Delta\mathbf{x}$ is the current perturbation in the state and $\Delta\mathbf{y}$ is the separation (including errors due to noise) between the measurement and the measured state. Recall that we have dropped the matrix form of K_{ij} as we are assuming the observable quantities are a subset of the states and not a function of several or all of them as would necessitate off diagonal elements. Given the (coupled) evolution of the state over time:

$$x_i(t+dt) = x_i(t) + dtf_i(\mathbf{x}) + K_i(y_i(t) - x_i(t)), \tag{3.14}$$

we note setting $K_i = Rm_i/dtRf_i$ captures the accuracy of both measurements and model for Gaussian noise in both model and measurements.

Obviously there is no requirement that these perturbations be the same - or even in the same direction. Should the system reach a point where the path is within the noise levels for both model and measurement, the evolution of $d\mathbf{X}/ds$ will be highly chaotic. Given this occurs when the path is within the margin of error this is a non-issue as the resulting attractor will likely be confined within those bounds. Monte Carlo methods, such as in [Quinn & Abarbanel, 2010], can be used to fully sample the distribution beyond this point.

The importance of this is not that we now have a meaning for the coupling strength, but that this suggests something fundamental about our ability to estimate states and parameters. The dynamics of d$\mathbf{X}$/ds has unstable directions if the path it is searching over contains any *local* instabilities that are not smoothed over by appropriate coupling. Because we are perturbing the entire path, we cannot resort to the state being on average regular as we could with forward evolution of the states

in time. Thus, even though the FHN model (Eq. 2.14) is globally regular, if the path contains local instabilities caused by the bifurcation between fixed point and limit cycle behavior, we will not be able to estimate the state of the system without measurements that tell us where and when this threshold is passed. Note that this is a qualitative argument, there doesn't appear to be any direct relation between the size of Rm/Rf and K and the ability to estimate states by finding extrema of the action.

The other side of this when the dynamics are always regular and there are no local instabilities - $\sum_j \left[\delta_{ij} + dtJ_{ij}(\mathbf{x}(t))\right] < 1$ for all $\mathbf{x}$(t), d$\mathbf{X}$/ds will evolve to the global minimum fixed point without the need for measurements. For regular systems with only occasional instabilities (such as spiking neurons) this suggests that the total number of measurements (what times how often) needed to estimate the full state of the system may be surprisingly low. The absence of measurements may prove a hindrance if the model parameters are not known a priori, as there is no way to estimate the state of a system over time if you do not know how it evolves

3.3 Parameter Estimation

The path integral method provides us with the means to estimate unmeasured states and parameters conditioned on measurements of some subset of the variables. While which states needs to be measured can be determined based on the synchronization of coupled systems discussed previously, this discussion has until now ignores that the model parameters are in general unknown.

Parameter estimation is a different problem from the synchronization in states. State synchronization is determined by whether there is convergence in state space towards a common trajectory over time. Synchronization in parameters requires not only synchronization in states but sufficient information from measurements to determine the set of parameters that would lead to the state being in a given region at a given time. For the Langevin equation on our distribution, this corresponds to a step of:

$$\frac{dp_i}{ds} = -\frac{\partial A_0}{\partial p_i} = \sum_{t=0}^{N}\sum_{j=1}^{D}\Big(\dot{x}_j(t) - f_j(\mathbf{x}(t),\mathbf{p})\Big)Rf_j\frac{\partial f_j(\mathbf{x}(t),\mathbf{p})}{\partial p_i} \tag{3.15}$$

The evolution of the parameters are sensitive to the errors in the model through differences in the local dynamics and to the state dependent sensitivity of those dynamics to each parameter. This is non-trivial as the dynamics themselves (and thus the errors in the dynamics) are a function of these parameters.

Naively we postulate that if our measurements are sufficient to synchronize model and data, those measurements are also sufficient for estimating model parameters given sufficient time. For the Lorenz 63 system:

$$\begin{aligned} \dot{x}_1 &= \sigma(x_2 - x_1) \\ \dot{x}_2 &= -x_2 + rx_1 - x_1x_3 \\ \dot{x}_3 &= -bx_3 + x_1x_2 \end{aligned} \tag{3.16}$$

we previously demonstrated (see 2.5.2) the ability to synchronize two *identical* oscillators with appropriate coupling to any one of the three dynamical variables - with optimal synchronization (as defined by the most negative largest CLE) occurring at around a coupling strength of $K_i = 10$ (Fig. 2.5) for coupling to x_2 and x_3 and 1000 for coupling to x_1. As was discussed earlier (see 3.2), K_i is roughly equivalent to $Rm_i/dtRf_i$. Without loss of generality, we can fix Rm at 1. For dt = 0.01, this suggests an Rf of 10 for the dynamical variables.

To test whether synchronization in state space implies the ability to estimate model parameters we use a variational method to minimize the action of the path integral formulation (Eq. 3.9). Estimates are made using the IPOPT optimization library with a python front end (see Ap. A) to generate the cost function and constraints from the equations of motion.[1]

For each of the three variables. we provide time series measurements of 10000 consecutive points at a 0.01 time step - corresponding to several dozen 'circuits' around the attractor. For measurements of x_1 estimates of model parameters are within a factor of $1e-4$ of their true values. Measurements of x_2 and x_3 are unable to

[1]It is important to note that IPOPT is not preforming a gradient descent on A_0, but rather a Newton search for extrema of A_0. Thus there is no guarantee that the minima it finds is the same minima as would be found using a gradient descent.

Table 3.1: Parameter estimates using time series measurements of each of the three state variables of the Lorenz 63 oscillator. Only measurements of x_1 yielded accurate parameter estimates. There did not appear to be a direct correspondence to the coupled synchronization - which would suggest estimates of parameters for coupling to each state. * indicates the parameter is at one of the variable bounds.

Meas.	σ	ρ	β
x_1	16.0	40.0	1.00
x_2	100*	100*	10*
x_3	26.1	26.1	1.2e-4
True	16	40	1

produce quality estimates of the parameter values. (Tab. 3.1) Surprisingly, parameter estimates using data from x_1 are within %10 for as few as 10 data points and within %1 with 100 - much less than one circuit around the attractor.

This indicates that the ability to synchronize states through coupling to a particular measurement is not sufficient for parameter estimates. This is not particularly surprising given the synchronization applies only to sufficiently similar models. A possible alternative is to look at the leverage provided by a particular state variable. x_1 appears in all three components of the equations of motion, while x_2 and x_3 only appear in two components. Future exploration of this question is warranted, but for now it is sufficient to say that our simulated experiment has demonstrated that measurements of x_1 are required to estimate parameters in the Lorenz 63 system.

For driven systems, such as the neuron models we will consider later, the measurements we can make of the system are a function of both how the model is driven and of the intrinsic dynamics of the system. Because accurate parameter estimates require a level of resolution in the underlying dynamics of the system and how the sensitivity of those dynamics to that parameter, we must insure that the stimulus provided to the system provides that resolution.

Given that the dynamics of the system is ultimately parameter dependent, there cannot be an ideal stimulus waveform for estimating parameters of a given model absent model constraints. While repeated experimentation on the real system with iterative changes to the stimulus may eventually result in a correct stimulus routine, in many cases these experiments are costly in terms of both materials and time.

Table 3.2: Estimated parameter values using the noisy measurements of x_1 in Fig. 3.1. 'Scale' refers to the scaling applied on the stimulus in Fig. 3.1. Weaker stimuli do not explore as much of the dynamical range as stronger stimuli. The quality of estimation is improved as more of the dynamical features are explored. Parameter bounds were chosen to be roughly an order of magnitude in each direction from the true values. ** lower bound for b*

Scale:	0	0.2	0.8	Act.
τ^{-1}	0.140	0.085	0.080	0.08
a	1.1	0.80	0.70	0.7
b	0.1*	0.62	0.79	0.8

The FHN system (see - 2.3):

$$\dot{x_1} = x_1 - \frac{x_1^3}{3} - x_2 + I_{ext} \tag{3.17}$$

$$\tau \dot{x_2} = x_1 + a - bx_1 \tag{3.18}$$

exhibits stimulus dependent bifurcations between Figure 3.1 (left) shows the simulated measured traces of 3 driven FHN neurons with a variably scaled, yet otherwise identical stimulus. $\sigma = 0.1$ Gaussian noise is added to measurements to simulate noise in the system. The minAzero program is used to estimate parameter values, as shown in Tab. 3.2, and all states at all times. The last state values and parameters are used as initial conditions to predict the model forward. Using the same stimulus after the estimation window, Figure 3.2 shows that a combination of different types of currents (Figure 3.1 right, t = 200 to 400) reveals that only the stimulus that fully explored the behavior of the system (the strong stimulus) was able to estimate parameter sets that can be used to predict future behavior. This procedure will be key for more complicated neuron models later in this thesis (see - 4.2).

3.4 Correspondence to the Extended Kalman Filter

A common method for estimating unmeasured states and parameters in nonlinear models is the use of what is called an Extended Kalman Filter.[Julier and Uhmann, 1997] Kalman filters are used to update the state and distribution of a

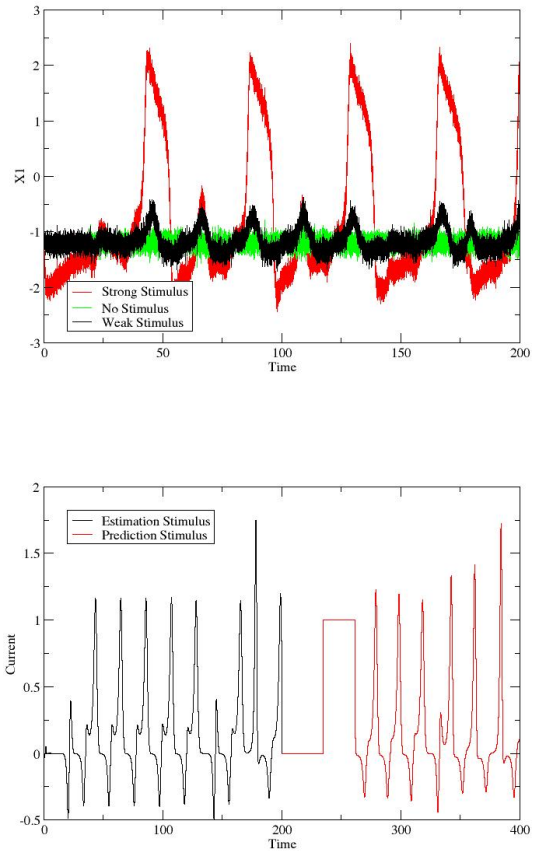

Figure 3.1: Plot of x_1 vs time for driven FHN neuron model with added $\sigma = 0.1$ Gaussian noise driven by a variably scaled stimuli. The driving stimulus during the estimation period (right, black) is scaled by a factor of 0.8 for the strong stimulus, and by 0.2 for the weak stimulus. These measurements are used to estimate the unmeasured x_2 state and unmeasured model parameters.

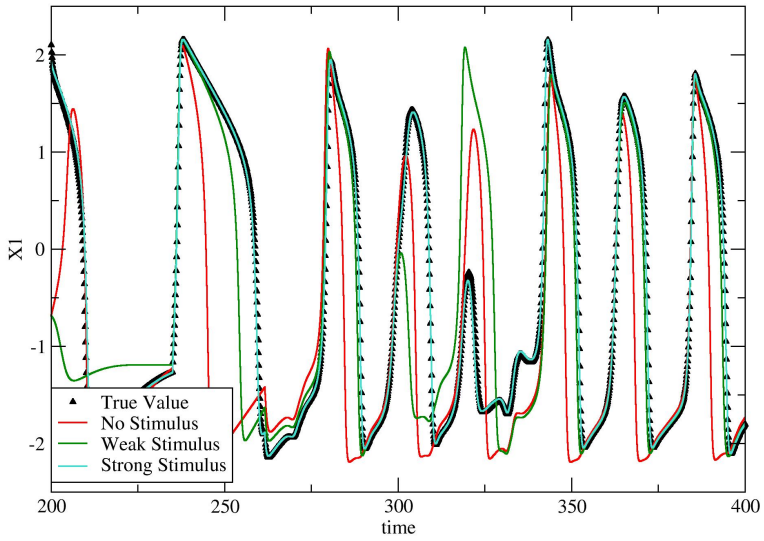

Figure 3.2: Plot of prediction of x_1 vs time for driven FHN neuron model. Initial conditions are determined by the final state of the estimation and the estimated parameters (Tab. 3.2) - though an initial period no stimulus eliminates any initial state information. The scaling in the estimation window is not applied to the stimulus in the prediction window (Fig. 3.1, right - red) to highlight the differences in behavior between the parameter sets. Only the parameter set estimated using the strong stimulus accurately tracks the true behavior.

system in the presence of underlying stochasticity and periodic noisy measurements of some function(s) of the dynamical variables:

$$\dot{\mathbf{x}}(t) = \quad \mathbf{f}(\mathbf{x}(t)) + \mathbf{w}_d(t) \tag{3.19}$$

$$\mathbf{y}(t) = \quad \mathbf{h}(\mathbf{x}(t)) + \mathbf{w}_m(t) \tag{3.20}$$

where $\mathbf{w}_d(t)$ and $\mathbf{w}_m(t)$ are the Gaussian distributed noise in the dynamics and measurements respectively. These noise terms are assumed to have respective matrix variances of Q_{ij} and R_{ij}. The expectation value ($\hat{\mathbf{x}}$) and the tensor variance about that value ($\mathbf{C}(t)$) evolve according to:

$$\dot{\hat{\mathbf{x}}}(t) = \quad \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{K}(t)\Big(\mathbf{y}(t) - \mathbf{h}(\hat{\mathbf{x}}(t))\Big) \tag{3.21}$$

$$\dot{\mathbf{C}}(t) = \quad \mathbf{J}(\hat{\mathbf{x}}(t))\mathbf{C}(t) + \mathbf{C}(t)\mathbf{J}(\hat{\mathbf{x}}(t))^T + \mathbf{Q}(t) \tag{3.22}$$

where $\mathbf{J}(\hat{\mathbf{x}})$ is the Jacobian evaluated at the current expected value of $\mathbf{x}$(t) and

$$\mathbf{K}(t) = \mathbf{P}(t)\mathbf{H}(t)^T\mathbf{R}(t)^{-1} \tag{3.23}$$

and

$$\mathbf{H}(t) = \left.\frac{\partial h_i}{\partial x_j}\right|_{\hat{\mathbf{x}}}. \tag{3.24}$$

The cost function as developed previously assumes a special case of this where Q and R are diagonal. While this linearization is essential for the Kalman filter, the path integral method could in principle use any twice differentiable distribution - including state dependent noise. Absent foreknowledge of the noise in the measurements or model, the choice of noise is fairly arbitrary. We choose uncorrelated Gaussian noise for the simplicity. For directly measured states, $h_i(\mathbf{x}(t)) = x_i(t)$, the matrix H will be diagonal. For Gaussian distributed measurements and model errors, this turns K(t) into R_m/dtR_f, as was noted earlier (3.2) functionally equivalent to the effective coupling in the unconstrained path integral formulation.

The major difference between Kalman filter methods and the path integral method is the actual problem they are solving. Kalman filters are set up such that

a distribution of states is evolved over time according to the dynamics of the system. The initial state distribution is 'stretched' by noise (and chaotic behavior) then 'squished' by measurements of the system. While this is excellent for predicting future behavior it is not ideal for estimating model parameters as the effects of a perturbation can only be treated with locally. [Carrassi and Vannitsem, 2011]

The path integral method has the advantage of combining the local uncertainty in state to state transitions with the global trajectory of the system. Instead of trying to solve $(\mathbf{x}(t+1), \mathbf{p}|\mathbf{x}(t), \mathbf{y}(t+1))$ we instead are solving for $(\mathbf{X}(T), \mathbf{p}|\mathbf{Y}(T))$. As time independent parameters affect the dynamical evolution of all states at all times - the value of that parameter will generally be better determined according to the entire time series of measurements rather than through localized transitions.

Chapter 4

Individual Neurons

Neurons are the nodes of neural networks, functioning as dynamical logic gates in the processing of information from stimuli into instructions for behavior. This section will provide a brief overview of the biology of neurons and the formulation of some common models describing their behavior. This will then be related to the development of experimental procedures to estimate the various model parameters conditioned on the types of measurements available from current experimental techniques. The procedure is shown first using twin experiments (see section 2.5.1) using simulated data and then extended to real experimental data from neurons in vitro.

4.1 Neuron Physiology

Neuron as a leaky cable

Control of complex organisms requires active measurements of the surrounding environment and formulation of a response in motor control in a time scale comparable to that of the environment. Absent some controlled forcing, instructions in the form of molecules are transferred throughout the cell through random Brownian motion. Assuming the molecules carrying the information are conserved we have Fick's law:

$$\mathbf{J}_{diff} = -D\nabla[C] \tag{4.1}$$

where [C] is the local concentration and J_{diff} is the flux of ions due to diffusion. For most molecules in water at room temperature (roughly the conditions inside a cell) D

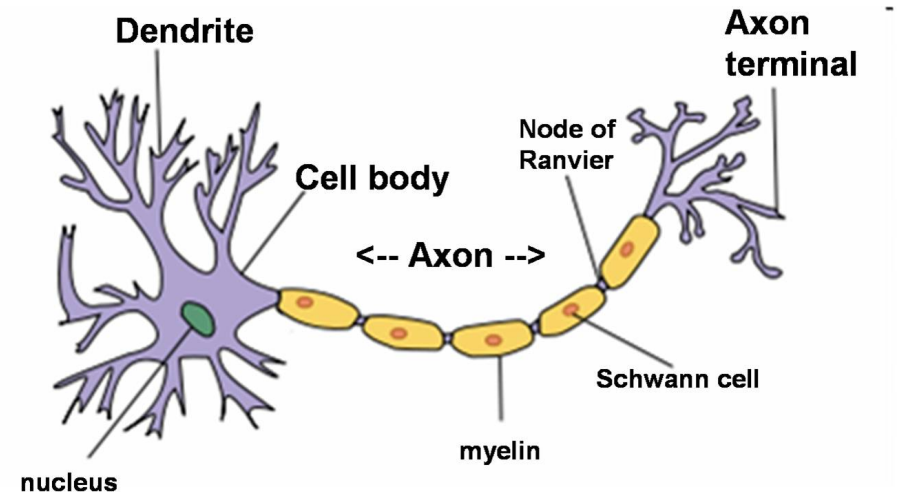

Figure 4.1: A qualitative diagram of a neuron showing the relative orientation of the soma (cell body), dendrites, and axon. Reproduced under CC-BY-SA-3.0 from wikipedia.org

is approximately 1e-5 cm^2/s. For a neuron that typically has length 1 mm (or more) this corresponds to a diffusive time scale of order tens of minutes.

Given the astronomical diffusion timescales at room temperature for objects of size order meters or centimeters, a faster, more directed means of transmitting information is needed. This speed is provided through the use of electrical signals over chemical signals. This electrical signal is created and propagated by the time varying concentrations of the various species of ions present in the body. Because they are charged, the effects of the addition of additional ions immediately effects all other charged ions through a change in the local electric potential. The rate at which information is transmitted is thus limited not by the speed at which particles move, but by the distance scale over which the field drops off and the time scale of the response. Because the membrane of the cell is essentially a leaky waveguide, this process is described by the cable equation:

$$\frac{1}{r_l}\frac{\partial^2 V}{\partial x^2} = c_m \frac{\partial V}{\partial t} + g_m V. \tag{4.2}$$

The longitudinal resistance, r_l is the resistance per unit length of signals propagating through the intracellular medium (cytoplasm), c_m and g_m are the per unit length capacitance and conductance of the cell membrane. V is the potential difference

between a point inside the cell and the exterior 'ground'[1]. These per unit length values are a function of the physical properties of the cytoplasm and membrane along with the geometry of the cell. For the roughly cylindrical geometry the longitudinal resistance scales inversely with increases in cross sectional area and proportionally with the length of the cylinder. The membrane resistance scales inversely with the circumference and cylinder length. To keep the scaling consistent we switch to the inverse value of conductance which scales proportionally with both circumference and length.

$$r_l = \frac{\rho_{cyt}}{\pi a^2} \tag{4.3}$$

$$g_m = \frac{2\pi a}{L\rho_m} \tag{4.4}$$

where a is the radius of the cylinder, d is the membrane thickness, and ρ_m and ρ_{cyt} are the resistivity (Ωcm) of the membrane and cytoplasm respectively. The distance over which these electrical effects act is described by the length constant of steady state solutions to the wave equation - which equals the cylinder length at which the membrane resistance is equal to the cytoplasm resistance:

$$\lambda = \sqrt{1/r_l g_m} \tag{4.5}$$

noting that r_l and g_m are both per unit length. The resistivity of cytoplasm is about 30 - 300 Ωcm. For a compartment of radius 1e-4 cm (area ≈ 3e-8 cm^2), r_l ranges from 1e10 to 1e11 Ω/cm. The membrane conductance is of order 1e-2 to 1 mS/cm^2, which corresponds to a per unit length conductance of 1e-6 to 1e-8 S/cm. This gives an order of magnitude for the length scale of about 1-100 μ m.

The time scale of the response to this signal is given by RC time constant of the membrane:

$$\tau = c_m/g_m \tag{4.6}$$

Which when combined with the characteristic length scale provides a characteristic velocity for the propagation of a signal down the neuron. For a typical neuron the per unit area capacitance is 1 $\mu F/cm^2$ and the membrane conductance is of order 1e-2 to 1 mS/cm^2, yielding a time constant of order 1 - 100 μs.

[1] While we assume that the exterior of the cell is at a uniform potential, this is only approximately true.

$$v = \lambda/\tau \tag{4.7}$$

Together this gives signal velocities of order meters per second - allowing for signals to be transmitted across large distances over short times.

Action potentials and the Hodgkin-Huxley Model

Electrical signals provide a fast way of transmitting information from one end of the cell to the other. However there is a catch to this process. While a locally applied stimulus will cause a signal to propagate down the cell, the signal will disperse over order 10 μ m due to the non-zero resistivity of the cytoplasm. Given that neurons are of order 1 mm in length or longer, there must be a means to regularly refresh the voltage signal over time.

The flow of each specie of ion across the cell membrane is equivalent to a battery connected to a variable resistor. For each ion the 'battery' is driven by differing concentrations of a particular specie of ion inside and outside the cell, leading to a bias in the diffusion. When combined with the ions carrying a charge, this bias creates an electrical potential. This potential is given by the Nernst potential.[2] [Johnston and Wu, 1995] and describes the state when the diffusive motion in one direction from a concentration gradient) is canceled by the electrical drift in the other from the potential caused by that concentration gradient. While diffusion may be ineffective for large distances, for the short distance across a cell membrane, the time scale of the motion of ions is of order 1 ns - effectively instant compared with the μs - s time scales prevalent in neuron behavior.[3]

$$E_k = \frac{RT}{qF} \ln \frac{[k_{out}]}{[k_{in}]}, \tag{4.8}$$

where R is the molar gas constant, T is the temperature, q is the charge of the ion (in units of the charge of the proton), F is the Faraday constant and $[k_{out,in}]$ are the molar concentrations of the ion inside and outside the cell. This 'battery' is attached

[2]Note that for many channels, multiple species of ions are exchanged simultaneously. In such cases the Goldman-Hodgkin-Katz equation is used. This form derives from the Nernst equation.

[3]The diffusion constant for each ion are of order 1e-5 cm^2/s and the membrane is about 1 nm thick.

to a 'resistor' corresponding the time dependent permeability of the cell membrane to a particular species of ion. This gives a formula for the current through the cell membrane of a particular ion as a function of time:

$$J_A(t) = g_k(t)(V(t) - E_k). \tag{4.9}$$

This, along with the conductance and capacitance of the membrane itself provide an RC circuit equation that yields the voltage response of the neuron:

$$C\frac{dV}{dt} = -\sum_k J_k(t) \tag{4.10}$$

where the sum on k is over the various types of ions that determine the membrane potential. These ions typically includes sodium and potassium, with calcium and chloride channels also common. This equation is typically expressed in terms of the per area behavior - C has units of pF/cm^2 and g(t) has units of mS/cm^2. While E_k is roughly constant for most ions ($|[\dot{k}]/[k]| << 1$), in many neurons the interior calcium concentration is low enough (and the exterior high enough) that the flow of ions into and out of the cell drastically changes the interior concentration. In such cases, the reversal potential is also time dependent. We do not consider such channels in this work.

The time dependences in g(t) come from the presence of specialized membrane proteins that stochastically undergo conformational changes that either open or close a pore for a particular type of ion. This leads to a modification of the current for a particular ion:

$$J_k = [\bar{g}_{ka}a(t) + g_{kp}](V - E_k) \tag{4.11}$$

where $\bar{g}_{ka}$ is the maximal conductance per unit area of the type of ion due to these membrane proteins, a(t) is the fraction of these channels that are open at a given time. g_{kp} is the passive conductance per unit area of the cell's lipid bilayer membrane. For simplicity all of these passive channels can be combined into a single 'leak' current:

$$J_L = \sum_k g_{kp}(V - E_k) = g_L(V - E_L). \tag{4.12}$$

Since the number of membrane proteins per unit area is fairly large (order 1e10/cm^2 for sodium and 1e9/cm^2 for potassium) [Hille, 2001], the stochastically

changing conformational state of each protein can be treated statistically in aggregate. Given the sodium and potassium membrane proteins contains each contain four gating sites, the channel becomes open only when each of the four gating particles are in the correct conformational state.[4] For the the potassium channel this means all 4 are open and the for the sodium channel, the 3 activating particles are open and the deactivating particle is not closed. Assuming the state of each gating particle is uncorrelated with the others, this means that the current for the active sodium and potassium channels are given by:

$$\begin{aligned} J_{Na^+} &= \bar{g}_{Na^+} m(t)^3 h(t)(V(t) - E_{Na^+}) \\ J_{K^+} &= \bar{g}_{K^+} n(t)^4 (V(t) - E_{K^+}) \end{aligned} \tag{4.13}$$

where m is the fraction of sodium activation particles in the open configuration, h is the fraction of sodium deactivation particles in the open configuration, and n is the fraction of potassium activation particles in the open configuration. In many cases a particular ion will be moderated by several types of membrane proteins. In these cases the time dependence will be a sum over different types of kinetics with maximal conductance proportional to the concentration of the respective protein in the membrane. The reversal potential is determined by the Nernst potential for that ion, and remains the same across different types of proteins.

The fraction of states in the open configuration (m, h, and n above) at a given time is given by Boltzmann statistics with the energy change between open and closed states a function of the membrane potential. Since the transitions are not instantaneous - and frequently slower than the changes in the voltage - these fractions are dynamical equations in time.

$$\dot{a} = \alpha(V)(1 - a) - \beta(V)a \tag{4.14}$$

$$\alpha(V) = a_\infty(V)/\tau_{open} \tag{4.15}$$

$$\beta(V) = \big(1 - a_\infty(V)\big)/\tau_{close} \tag{4.16}$$

[4] In the original work on the voltage behavior of neurons, the structure of the various gating proteins were not known - the choices of exponents were qualitative fits. While current knowledge of the structure of these proteins appears to a validate these choices - the lack of knowledge about state correlation means the exponent is still somewhat of a qualitative property.

with τ_{open} and τ_{close} the time constants for changing to the open and closed states respectively. This is qualitatively equivalent to:

$$\dot{a} = \frac{a_\infty(V) - a}{\tau_a(V)} \tag{4.17}$$

where $a \in \{m, h, n\}$. The exact energy function of voltage for each conformational state cannot be found from first principles - but are instead known from qualitative fits based on work started by Hodgkin and Huxley ([Hodgkin and Huxley, 1952]) on the squid giant axon.[5] The behavior appears to obey a smooth sigmoid function - we choose to use a tanh formulation.

$$a_\infty(V) = \frac{1}{2}\left(1 + \tanh\left[\frac{V - V_{\frac{1}{2}a}}{dV_a}\right]\right) \tag{4.18}$$

The time constants are similarly qualitatively fit with:

$$\tau_a(V) = \tau_1 + \tau_2\left(1 - \tanh^2\left[\frac{V - V_{\frac{1}{2}a\tau}}{dV_{a\tau}}\right]\right) \tag{4.19}$$

Where $V_{\frac{1}{2}a\tau}$ and $dV_a\tau$ are frequently assumed to be equal to $V_{\frac{1}{2}a}$ and dV_a respectively based on qualitative observations about the sensitivity of the dynamics of the model to relative differences in those values. [Toth, et al., 2011]

Equations 4.14, 4.12, and 4.17 together form what we will refer to as the 'simple' Hodgkin Huxley model:

$$C\dot{V} = \bar{g}_{Na}m^3h(E_{Na} - V) + \bar{g}_K n^4(E_K - V) + g_L(E_L - V) + I_{ext}(t)/A \tag{4.20}$$

$$\dot{a} = \frac{a_\infty(V) - a}{\tau_a(V)} \tag{4.21}$$

where $a \in \{m, h, n\}$ $I_{ext}(t)$ is an external driving current applied through the patch of cell. As this current is not applied uniformly it must be scaled by the surface area of the local compartment.

The actual behavior of the Hodgkin Huxley model depends on the relative values of the various time constants, maximal conductances, and activation levels. While much has been written about the bifurcation structure of this simple version

Table 4.1: Parameters in the NaKL HH Model: Na, K, Leak Currents. While the underlying properties of the neuron may vary substantially the values here correspond well to what has been observed in experiments on individual ion channels in isolation.

Name	Value	Name	Value
C_m	1.0 $\mu F/cm^2$	Vh	-60.0 mV
g_{Na}	120.0 mS/cm^2	dVh	-15.0 mV
E_{Na}	50.0 mV	t_{h0}	1.0 ms
g_K	20.0 mS/cm^2	t_{h1}	7.0 ms
E_K	-77.0 mV	Vht	-60.0 mV
g_L	0.3 mS/cm^2	dVht	-15.0 mV
E_L	-54.4 mV	Vn	-55.0 mV
Vm	-40.0 mV	dVn	30.0 mV
dVm	15.0 mV	t_{n0}	1.0 ms
t_{m0}	0.1 ms	t_{n1}	5.0 ms
t_{m1}	0.4 ms	Vnt	-55.0 mV
Vmt	-40.0 mV	dVnt	30.0 mV
dVmt	15.0 mV	I_{DC}	-8.0 pA/cm^2

of the model [**?**] and the various types of behaviors allowed by this model, we choose to use biophysically plausible values. (Table 4.1)

For this parameter set, the neuron will undergo what is referred to as a spike under certain stimuli. A positive influx of current will depolarize the neuron from its resting state near the leak reversal potential (E_l). As the voltage rises the sodium activation variable (m) increases and deactivation variable (h) decreases as the voltage approaches and passes their respective midpoints (V_m and V_h). Because of the relative time scales, the activation variable will open the channel faster than the deactivation variable can close it (0.1 ms vs 1 ms). This provides a window where m^3h is non zero and given the very large maximal conductance for sodium drives the neuron very quickly towards the sodium reversal potential at +50 mV.

While this is happening the potassium channel (n) has been activating. Because of its shallower voltage response, slow time constants, and 4th power scaling; the potassium current only becomes active at an appreciable level after the neuron is depolarized. Since the reversal potential of this channel is at -77 mV this current repolarizes the neuron by driving the voltage back down below the level the sodium current activates. The sodium channel does not reactivate during this reset - despite

[5]Not to be confused with axons of a giant squid.

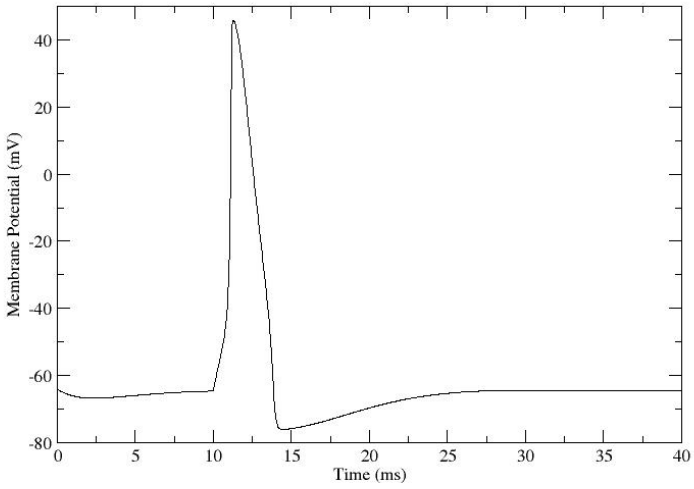

Figure 4.2: A voltage spike in response to a short pulse of positive current at t = 10 ms.

passing through the same activation region because the sodium deactivation particle remains active. Because the voltage responds faster than the gating particles, the recovery will overshoot the resting state of the neuron leading to below resting refractory period. The voltage response to a short pulse of current is shown in figure 4.2 and gating particles and the resulting currents in figure 4.3.

4.2 Single Neuron Twin Experiments

Estimation of the unmeasured parts of any model requires properly leveraging the data available from measurements. For individual neuron models the data available typically comes from patch clamp recordings [Hamill, et al., 1981] that are able to measure voltage across a small patch of the cell membrane (typically in the soma) while applying a known current waveform to stimulate activity. While experiments have been done on the activity of individual membrane proteins (see examples in [Hille, 2001]) there is no current tool that can measure the fraction of all local gating particles that are in a given configuration. Likewise, while certain characteristics of the neuron - represented through model parameters, can be known from other exper-

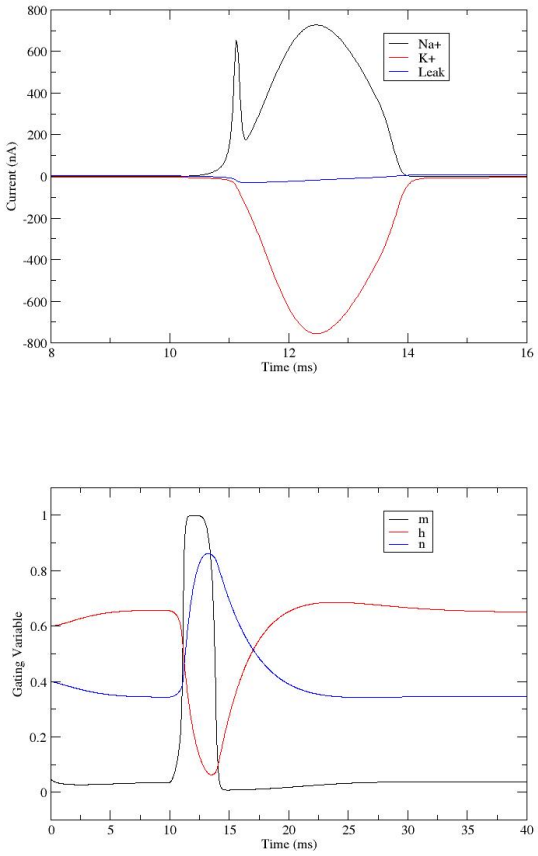

Figure 4.3: Anatomy of a spike - sodium, potassium and leak currents as a function of time during and an action potential (right) and m, h, and n gating variables (left). Recall that the ionic currents scale with distance from their respective reversal potential.

iments (such as the area of the cell and the capacitance per unit area) most values must be inferred through a combination of available measurements and assumptions about the structure of the model.

Because of the difficulty and cost of preforming experiments on live neurons - particularly for in-vitro recordings that require killing the animal, simulated experiments are used to develop routines that can later be used to determine the properties of real neurons.

The procedure for estimation is twofold - as described in the previous chapter. Estimating the full state of the system for a given parameter set requires that there be sufficient information from the measurements to synchronize the model with the measured system. Estimating the parameters requires that the measured trajectory be sufficiently sensitive to the parameter values - which is a function of the choice of driving stimuli. Since in a twin experiment we have access to all states and parameters of a model in a simulated system it is feasible to check whether a given set of measurements will result in both synchronization in state space and well fit parameters. The twin experiment is thus a useful tool in determining the types of measurement and stimuli routines that should be used to estimate the state of a real system.

In the most idealized case where there is full knowledge of both the form of the model and its parameter values it must be checked whether the voltage measurements alone are sufficient to estimate the full state of the system. Since the Hodgkin Huxley model with only sodium and potassium channels is regular for biologically realistic currents, all possible trajectories will converge[6] given the same driving current. (see Fig. 4.4) This suggests that no measurements are needed to determine the state of the system at some later time - the model neuron undergoes generalized synchronization with the current waveform. We will return to this apparent property later when I turn our focus to characterizing the behavior of synapses and networks of neurons, for now it is sufficient to note that because the behavior of the Hodgkin Huxley model with the parameters chosen is globally regular, it passes the necessary condition[7] that

[6]For tonicly spiking neurons, the trajectories will differ by a phase shift.

[7]I have not come across a system where the former is not a necessary condition for the latter. There is no mathematical proof, to my knowledge that this is the case.

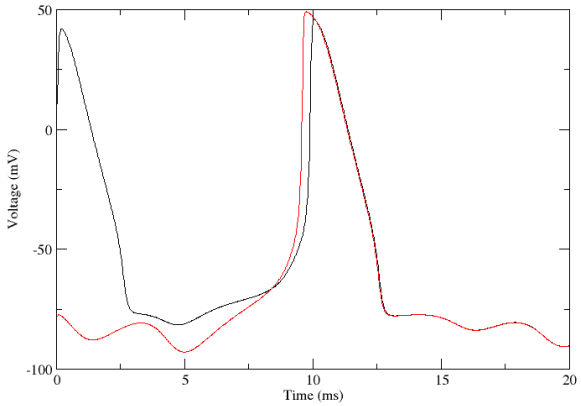

Figure 4.4: Voltage trajectory of two identical Hodgkin Huxley models given identical stimulus. The convergence in trajectories is a good indicator that the system is regular - no positive conditional Lyapunov exponents. Note that the synchronization is not local. During spikes, the trajectories are not converging, indicating that while on average (globally) the system is regular, this does not guarantee that the the system is locally regular.

voltage can potentially be used to estimate the full state of the system.

With only limited knowledge of the parameters - such as order of magnitude and whether it is positive or negative, we must devise a stimulus routine that has sufficient leverage over each parameter to simultaneously determine all of them. At the general, qualitative level, this means that estimating parameters requires exploring the areas of state space where those parameters are used. For a Hodgkin Huxley model this means exploring both the linear RC sub-threshold behavior and the active spiking behavior. Achieving this will thus require the application of both hyper-polarizing (negative) and depolarizing (positive) waveforms.

A common stimulating procedure in the literature is to use a series of step currents, both positive and negative, to stimulate the required hyper-polarizing and depolarizing currents.[Daou, et al, 2013] While useful for finding static properties of neurons (such as the passive RC time constant) these types of currents do not have sufficient complexity to estimate kinetic properties consistently. We can demonstrate this failure by comparing the estimates of the simple Hodgkin-Huxley model (Eq. 4.20) driven by a series of steps with a the results from the same neuron driven by a more

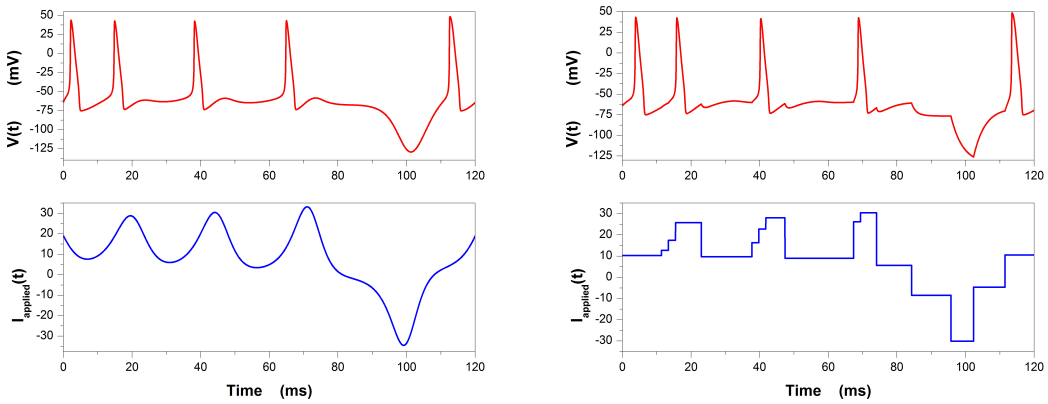

Figure 4.5: Plots of current vs time and voltage vs time for a simulated Hodgkin-Huxley neuron stimulated with step currents (Right) and a waveform from the Lorenz 63 model (Left). The steps were created by averaging over the Lorenz current between inflection points to create qualitatively similar activity patterns.

complex, smooth waveform.

To demonstrate this we consider a the response to a waveform from the Lorenz 63 system (see chapter 2.1) and the response to a series of steps that is roughly a histogram of that same current (Fig. 4.5). The two current produce qualitatively the same features at the same time so in principle should produce similar estimates. However when we run the estimation procedure using otherwise identical conditions - no model or measurements noise, Rm = 1 Rf = {10,1e5,1e5,1e5}, we find that only the smooth current produces good parameter fits (Tab. 4.2) in fact the estimation procedure consistently fails to converge to a answer, suggesting an unstable search.

This does not conclude that step currents are incapable of estimating the full properties of a neuron. It does suggest that complex smooth waveforms may be better at estimating many of the parameter values. More work is needed to develop a picture of the cost function landscape for various stimulus protocols. Given the number of possible stimuli are basically infinite, this is no easy task.

Unlike this ideal case the model will generally have inherent stochasticity to it and the measurements will generally be noisy. To simulate these errors we add a stochastic forcing in the form of an added $\sigma = 5$ pA Gaussian current to simulate noise in the system and a $\sigma = 1$ mV added Gaussian noise to measured state to simulate noise in the voltage probe. The stimulus current (absent noise) and the simulated

Table 4.2: Parameter estimates for a simulated experiment on a Hodgkin Huxley neuron stimulated with a series of step currents or a current generated from the Lorenz 63 model. For the same time window, the Lorenz current provides substantially better resolution of parameter values despite having the same number of spikes and same general shape. *Because of a multiplicative degeneracy between the capacitance, maximal conductances and surface area, we fix the capacitance at 1 pF/cm^2. ^{u} *at upper bound,* ^{l} *at lower bound*

Name (Units)	Actual	Steps	Lorenz	Name (Units)	Actual	Steps	Lorenz
C ($\mu F/cm^2$)	1	1*	1*	Area (cm^2)	0.8	1.45	0.80
g_{Na} (mS/cm^2)	120	200^u	119.9	E_{Na} (mV)	50	60^u	50.0
g_K (mS/cm^2)	20	40^u	19.9	E_K (mV)	-77	-79.1	-77.0
g_L (mS/cm^2)	0.3	0^l	0.30	E_L (mV)	-54	-55.3	-54.1
V_m (mV)	-40	-31.4	-40.0	dV_m^{-1} (mV^{-1})	0.0667	0.025	0.0667
τ_{m1} (ms)	0.1	0.05^l	0.10	τ_{m2} (ms)	0.4	0.1^l	0.40
V_h (mV)	-60	-40^u	-60.0	dV_h^{-1} (mV^{-1})	-0.0667	-0.0353	-0.0666
τ_{h1} (ms)	1	0.64	1.0	τ_{h2} (ms)	7	1.17	7.0
V_n (mV)	-55	-70^l	-55.0	dV_n^{-1} (mV^{-1})	0.0333	0.011	0.0334
τ_{n1} (ms)	1	5^u	1.0	τ_{n2} (ms)	5	12^u	5.0

measurements are shown in Fig. 4.6 (Black). Using the same procedure as before, we estimate the unmeasured states and parameters (Red). Despite the added noise, the parameters are estimated to within a reasonable margin of the true values as shown in Tab. 4.3. Since in a real experiment we can't simply compare the parameters to the true values, we also take the estimated final state and parameters as initial conditions for the model and integrate it forward to predict future behavior (Blue). This predicted voltage behavior is in good agreement with the observed behavior - a good indication of a quality fit using only the information available under a real experiment.

4.3 Model Estimates on Data from Real Neurons

The procedure used to estimate unmeasured states and parameters is ultimately intended for use on data measured obtained from real neurons. The results of these experiments will then be used to estimate properties of larger more complex structures in the brain of interest. More specifically:

- Characterize the types and respective dynamics of the neurons in the network

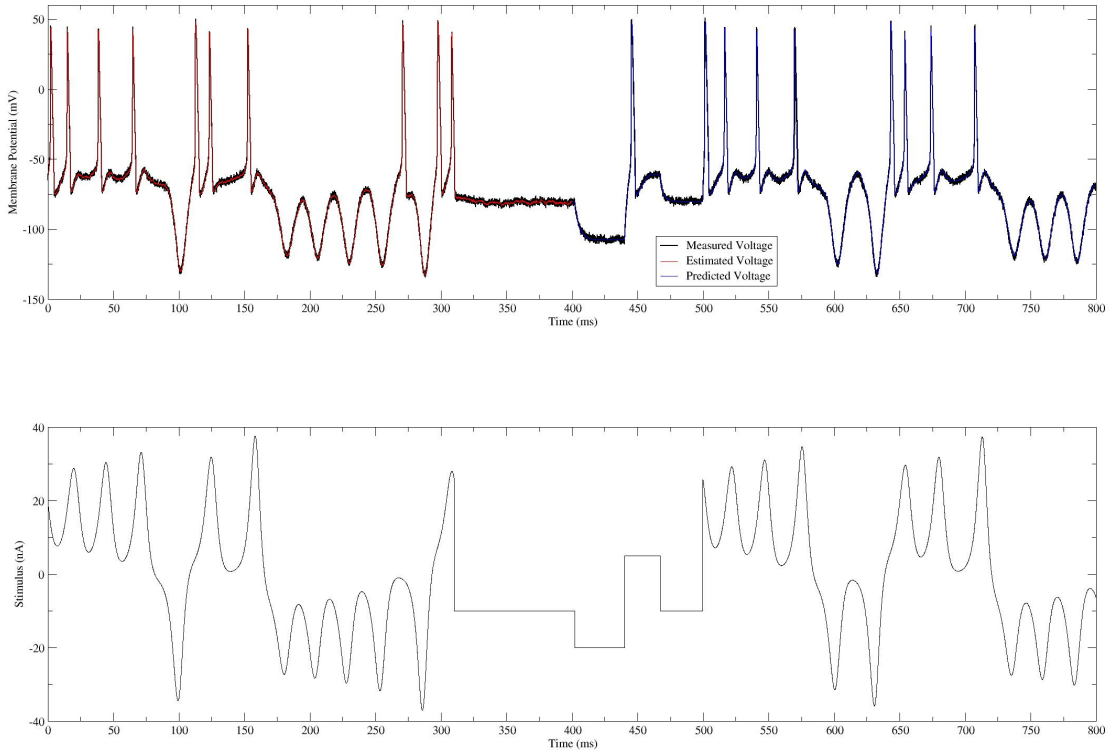

Figure 4.6: Plots of current vs time (left) and measured voltage vs time (right) for a simulated Hodgkin-Huxley neuron stimulated with a complex waveform from the Lorenz 63 system. The system has model errors simulated through the addition of a 1 mV/ms RMS Gaussian noise term in the dynamics. Errors in the measurement system are represented by a 1 mV RMS Gaussian noise added to the voltage state. The extra variability in both sub threshold and spiking regions provides sufficient resolution to estimate model parameters.

Table 4.3: Parameter estimates for a simulated experiment on a noisy Hodgkin Huxley neuron stimulated with a Lorenz current with added steps. The model was forced by an additional, unmeasured 5 pA Gaussian current to simulate noise in the system. An additive 1 mV Gaussian noise was added to the measured state to simulate noise the voltage probe. Despite the substantial noise, Because of a multiplicative degeneracy between the capacitance, maximal conductances and surface area, we fix the capacitance at 1 pF/cm^2.

Name (Units)	Actual	Estimate	Name (Units)	Actual	Estimate
C ($\mu F/cm^2$)	1	1*	Area (cm^2)	0.8	0.794
g_{Na} (mS/cm^2)	120	140.7	E_{Na} (mV)	50	49.7
g_K (mS/cm^2)	20	16.9	E_K (mV)	-77	-77.2
g_L (mS/cm^2)	0.3	0.293	E_L (mV)	-54	-53.4
V_m (mV)	-40	-39.6	dV_m^{-1} (mV^{-1})	0.0667	0.0667
τ_{m1} (ms)	0.1	0.103	τ_{m2} (ms)	0.4	0.389
V_h (mV)	-60	-61.5	dV_h^{-1} (mV^{-1})	-0.0667	-0.0615
τ_{h1} (ms)	1	0.934	τ_{h2} (ms)	7	6.28
V_n (mV)	-55	-55.8	dV_n^{-1} (mV^{-1})	0.0333	0.0344
τ_{n1} (ms)	1	0.889	τ_{n2} (ms)	5	5.36

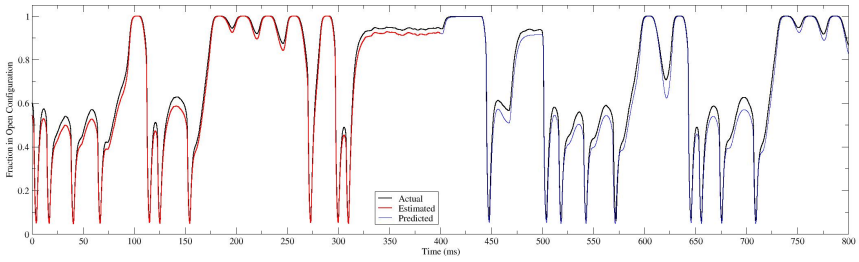

Figure 4.7: Plot of the unmeasured sodium inactivation gating particle - h(t) for a simulated experiment on an isolated neuron driven by a complex stimulating waveform. (Fig. 4.6) Estimates are made conditioned on regular (50 khz) noisy voltage measurements.

in isolation as described in this chapter.

- Use developed neuron models as a filter to characterize intranetwork and external currents present in the network such as synapses. (See Chap. 5)
- Use neuron and synaptic models to build a model of network behavior. (See Chap. 6)

The brain of interest for our group has recently been that of song birds, primarily the zebra finch. In particular we are interested in the behavior of a particular region of that brain: the HVC[8]. The HVC is a region of the avian brain involved with the learning and production of song. It projects (has neurons with axons that go to another region) to the anterior forebrain pathway - believed to be responsible for learning, and the posterior descending pathway which ultimately projects to muscle control in the syrinx.[9] [Nottenbohm, 2005] However, for now we are not interested in the full song production network as we need to first characterize the neurons and synapses in the HVC subnetwork.

To create accurate models of the behavior of individual neurons, those neurons should ideally be isolated from the rest of the network - ensuring that only the currents we provide are stimulating the observed responses. This isolation is achieved through the use of in vitro recordings in slice. Without going into the gory details[10] a slice from the network of interest is taken from the bird and the cells present are kept alive through a nutrient bath. Much like a sub-critical mass of plutonium, the reduced number of cells in the resulting slice cannot sustain continuous bursts of activity - resulting in all cells generally staying at their resting states. While some synaptic connections are still physically present, they are not activated. The physical removal of the subnetwork from the rest of the brain further prevents any out of network stimulus. Note that in some networks specialized pacemaker neurons will exhibit continuous activity in the absence of stimuli. [Ramirez, 2004] In such cases drugs may be needed to suppress the synaptic connections between cells. While these drugs

[8] originally *hyperstriatum ventrale, pars caudalis*, then high vocal center, now just HVC

[9] A vocal cord equivalent

[10] The bird is killed, drained of blood then decapitated. The skull is then cut open and the brain removed. The section of the brain of interest is then cut into thin slices and placed in a cooled, nutrient filled, oxygenated bath to keep the individual cells alive as long as possible.

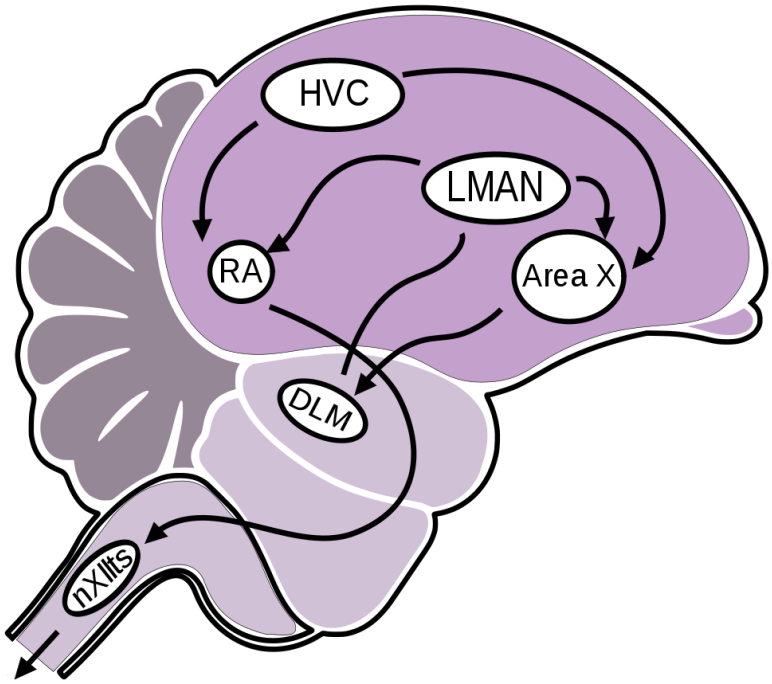

Figure 4.8: Rough network structure of the song related pathways in songbirds. HVC projects to the song production pathway via RA and to the song learning pathway via Area X. HVC receives recurrent input from the thalamic nucleus (motor function, not shown) and the auditory system (not shown). Original graphic from [Nottenbohm, 2005]. Reproduced under creative commons 2.5 license.

can be applied to be certain the remaining neurons receive no input we believe such methods should be avoided when possible due to possible unintended side effects on neuron behavior. However, HVC does not have these types of cells.

Once the network of interest has been separated, a neuron is identified for . The incentive for the slice preparation is not simply to isolate the network; in addition to isolating the neuron from rest of the network the thin slice allows for the visual identification of the neurons we are probing. The visual identification make application of the electrode easier and allows for the differing types of morphology to be identified prior to recording. Fig. 4.9 shows a typical single cell patch on an interneuron in a slice from the HVC of a zebra finch. [Daou, et al, 2013]

The electrode itself is a needle that uses suction to 'patch' into the surface of the membrane - effectively becoming part of the neuron membrane [Hamill, et al., 1981]. The electrode alternates between injecting a designer current through a saline solution and recording the voltage across the membrane at 20-50 kHz - much faster than any component of neural behavior. The 0.02 ms time step of our twin experiments corresponds to this sampling rate.

The procedure for estimating model parameters is then the same as for the simulated neurons, with the caveat that the simple Hodgkin-Huxley model with sodium, potassium and leak currents will be insufficient to describe the behavior of a neuron that is known to have types of channels beyond the minimalist spiking neuron model we have been working with. To that end, the model is expanded to include a few more channels:

$$-C\frac{dV}{dt} = J_{Na} + J_{NaP} + J_{K1} + J_{K2} + J_{K2} + J_{CaH} + J_{CaL} + J_C + J_{Leak} \qquad (4.22)$$

J_{Na}, J_{A1}, and J_{Leak} are the familiar sodium, potassium and leak currents, respectively, from the simple Hodgkin Huxley. J_{NaP} is a persistent sodium channel. J_{A2} and J_{K2} are other potassium channels. J_{CaH} and J_{CaL} are high and low threshold activated calcium channels, respectively. J_C is a calcium activated potassium channel and J_h is a hyperpolarization activated cation current.

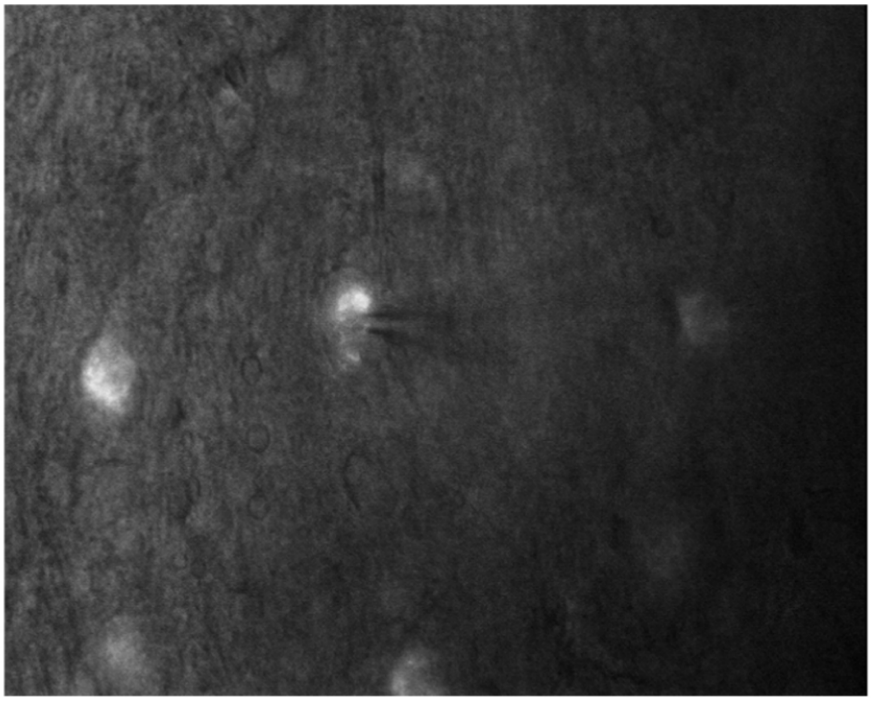

Figure 4.9: Image of patch clamp on interneuron in HVC. The probe is the dark lines at center. The light patches are the soma of neurons that have been illuminated using a florescent dye. Reproduced by permission of author from [Daou, et al, 2013]

Excepting the two calcium channels, the form of each of these channels is the familiar:

$$J_x = g_x a^i b^j (E_x - V) \tag{4.23}$$

where a is the activating variable (like m or n) and b is the inactivation variable (like h). The kinetics of a and b are of the form:

$$\dot{a} = \alpha_a(V)[1 - a] - \beta(V)a \tag{4.24}$$

with tanh formulations for the sigmoidal functions α and β.

Because the calcium concentration inside the cell is many orders of magnitude lower than outside the cell the linearized Nernst potential we used for the other channels no longer applies.

$$J_{Cax} = a^i b^j \left(\frac{g_{out} - g_{in} \exp(V/V_T)}{\exp(V/V_T) - 1} \right) V \tag{4.25}$$

where V_T is the thermal voltage for Ca^{2+} ($k_B T/(2e) \approx 13$ mV) g_{out} and g_{in} are proportional to the interior and exterior concentrations - the difference in concentration results separates the flow inward with the flow outward. Note that this equation has a removable singularity at V = 0. When this equation or its derivatives are dealt with numerically a Taylor expansion should be used to prevent potential issues.

Of course not all of these currents are certain to be present in the cell - nor is there a guarantee that there are additional currents that are present that are not included. Ideally a current that is not present will be estimated to have a maximal conductance (g_x) of 0. Generally we want to make the model as expansive as possible without being so expansive that the model becomes over fit.

However there is no easy way to determine what this point is. For twin experiments (see **??** we have full knowledge of the unmeasured states and parameters to compare with - even if we are not explicitly presenting that information to the estimation procedure. For experiments on real neurons the ability to determine the quality of the resulting model are more limited. Qualitatively, we can examine the estimated values of various parameters to check that they make physical sense. Quantitatively, we can use the estimated final state and parameters to estimate future behavior of the neuron then compare that with the observed future behavior.

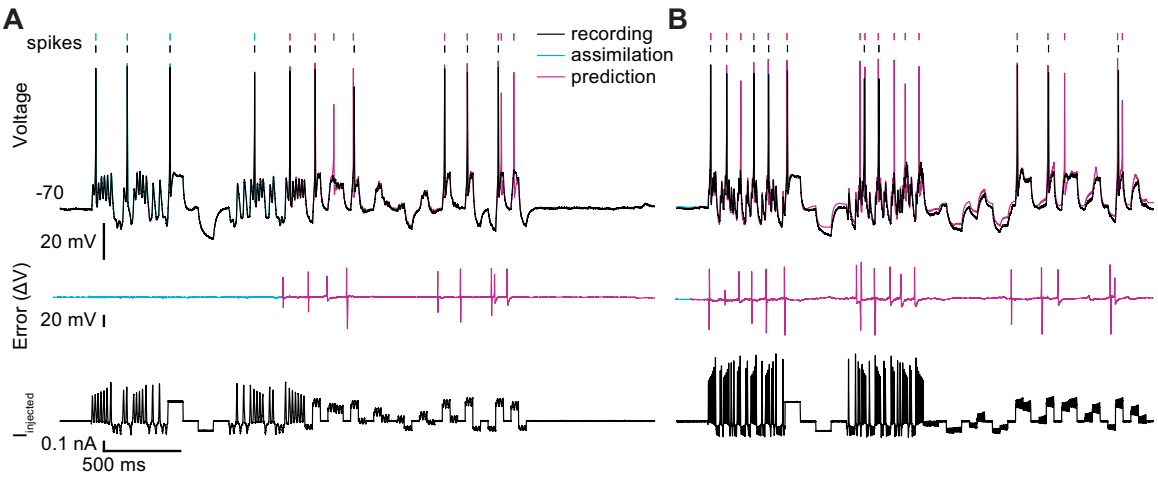

Figure 4.10: Plot of voltage vs time and stimulating current vs time for model estimation on an HVC interneuron recorded in vitro. The model is fit using the blue region in A, then used to predict future behavior from both the same data window (left) and from a unconnected data window (right). Some variation between time epochs is expected due to the changing conditions of the cell. Reproduced by permission from [Meliza, et al., 2014].

This predictability has been our primary metric to this point. Fig. 4.10 shows a period of x ms (y data points) used to estimate states and parameters of the expanded Hodgkin Huxley model. The final states and parameters of that model are then used to predict future behavior (pink). The model can also be used to predict the behavior of a novel stimulus without knowledge of initial conditions when provided a short (100 ms) window of to assimilate initial conditions[11]. The model does an excellent job at predicting the sub-threshold behavior, but there are some discrepancies between the spike shape, timing, and number in bursting periods.[Meliza, et al., 2014]

Acknowledgement

Figure 4.10 was reproduced by permission of authors from Meliza, C. D., M. Kostuk, H. Huang, A. Nogaret, D. Margoliash,and H. D. I. Abarbanel, "Estimating parameters and predicting membrane voltages with conductance-based neuron models" *Biol. Cybernetics*, **108** 495-516 (2014)

[11]Because some of the channels are slow or activated in the sub-threshold region, we cannot assume the system is at the resting state fixed point - or even that it has one.

Chapter 5

Characterizing Stimuli

Neurons are designed to receive stimulus from other neurons, integrate that stimulus according to rules defined by the types and distributions of active ionic currents then fire a spike that propagates down the cells axon to communicate its response with other neurons. The propagating voltage signal triggers the activation of specialized junctions between neighboring cells called synapses that can either depolarize (excite) or hyperpolarize (inhibit) the adjoining neuron.

Since the stimulus acting on neurons in a network are generally caused by these connections and not the experimentalist's needle, we must develop means for characterizing these connections. Unfortunately, the synapses themselves are tiny μm preventing direct experiments to determine their behavior. This chapter will describe the use of previously characterized neurons for which we have accurate models - as a filter to determine properties about the active synaptic connections.

5.1 Currents from Voltage - Using the Neuron as a Filter

Probing the behavior of synapses is difficult due to the size and diffuse scope of synaptic connections. Individual synapses are of order 1 μm in size - putting them below the ability to be resolved short of electron microscopy.[Gray, 1959] Further the connection between one neuron and another typically consists of many synapses at the many points at which the axon terminal of the pre-synaptic neuron and the

dendrites of the post-synaptic neuron come into contact.[Cowan, Südhof, & Stevens, 2001] Lastly even noting the presence of connections between the two cells is not sufficient information to determine the strength of the connection between the cells - which is determined by the density of receptor proteins in the synaptic cleft. These features make directly characterizing the behavior of synaptic connections difficult.

Since direct measurements of individual synaptic activity are unavailable - we must instead rely on the effects these connections cause in the measurable regions[1] through the currents that originate in the synapses - dynamically propagate through the synapse then cause a depolarization or hyper-polarization depending on whether the synapse was excitatory (AMPAergic) or inhibitory (GABAergic) respectively.

Probing synaptic activity can be accomplished using the measurements available from voltage recordings in the more accessible soma. Typically this procedure is done under voltage clamp [Zhang & Trussell, 1994, Destexhe, Mainen, & Sejnowski, 1994] - yielding excellent results about both the kinetics and maximal conductance of individual synapses. For networks of neurons, voltage clamping a single neuron effectively removes that neuron from the network as holding it at or near a specific voltage will prevent spiking activity - destroying its ability to respond by firing spikes.

Ideally, synaptic currents should be able to be inferred based on the response to recordings under a passive current clamp - as in the single neuron experiments except with $I_{app}(t) = 0$. In principle, the stimulating current - absent some model that couples it to the activity elsewhere in the system - is simply an additional dynamical state for which little is known about its evolution.

The little that is known is that in part because of filtering in the dendrites and the time scales of individual synaptic currents, the waveform of total synaptic activity is unlikely to include any high frequency components. Because of this we can use a trick in the path integral formalism to create a continuity and smoothness constraint by adding a cost to concavity.

$$\frac{d^2 I_{ext}(t)}{dt^2} \approx 0 \tag{5.1}$$

[1] it would be interesting to look at the activity in the dendritic compartment in response to synaptic activity. I was under the impression that such recordings were not possible - but apparently Michael Long is able to make them. I feel I am throwing a lot of potentially important biology under the rug by neglecting activity in the dendritic compartments.

with coupled first order ODE's for the current and its time derivative. The Rf/dt^2 term for these dynamics are chosen somewhat arbitrarily - but values around the value of Rm for the second derivative equation and several orders of magnitude larger for the (defined) dynamics of the first derivative equation have been found to work well at properly smoothing the current estimation. This smoothing will not greatly effect the estimation of real currents acting on the neuron as these tend to be relatively low in frequency compared with the sampling rate and the natural low pass filtering of the neuron membrane.

5.1.1 Twin Experiments

As a demonstration we consider the twin experiment used for the single compartment Hodgkin Huxley model considered previously. We use the fitting that was previously preformed as in Fig. 4.6 and fix the parameter values to the estimated values in Tab. 4.3. The *actual* values, that are used to generate the data, remain the same as before so the model will be as accurate as we could determine but not identical to the true system. Since state values are dependent on the unknown prior activity and unknown stimulus the unmeasured state values are also unknown.

Using simulated noisy measurements of the voltage ($\sigma = 1$ mV) we attempt to estimate the unknown stimulus and all unmeasured states using the variational method. The results of this, including the previously described continuity constraint are shown in figure 5.1. This demonstrates that the neuron stimulus can be recovered despite the presence of spiking behavior.

Implicitly, this procedure is a search for the low frequency components of the remainder of $dV/dt - f_V$, any errors in model will express themselves in this inverse problem. In this case, we had excellent prediction of the model parameters governing subthreshold behavior (g_L, E_L, Area) but incorrect (yet sufficiently accurate for prediction) values for some of the parameters that define the spiking behavior, notably g_{Na} and g_K. This suggests that this procedure could be used as a tool to quantify model errors beyond relying solely on the model's ability to predict future behavior as was developed and used in prior work. [Toth, et al., 2011, Meliza, et al.,

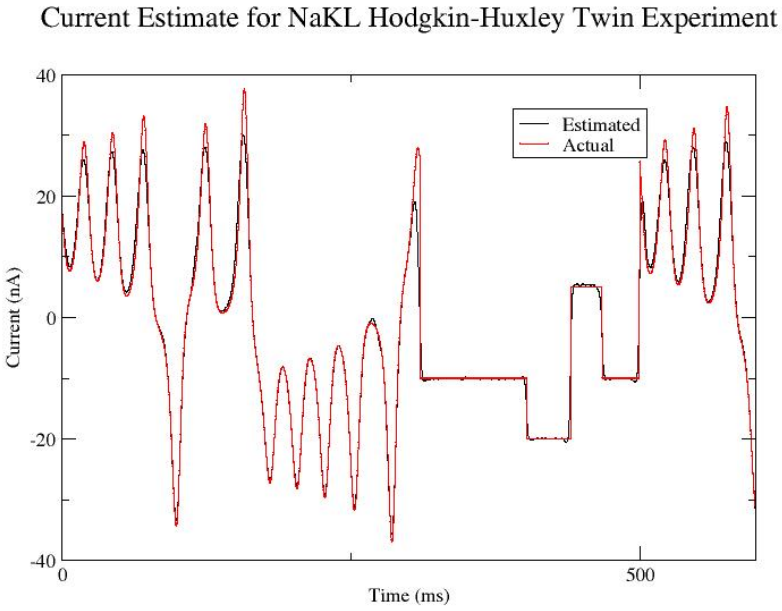

Figure 5.1: Estimates of current vs time based on noisy time series voltage measurements of a simulated identified neuron. Neuron parameter values are set according to the optimization procedure done with known stimulus in 4.2. A continuity/smoothness constraint is added by the inclusion of an additional cost term to the curvature of the current. This term is essential to keep the otherwise unknown dynamics of the current from attempting to track the measurement noise. Rm = 1, Rf = 10 for V and 1e5 for m,h,n. The smoothing Rf were chosen to be 100 and 10 for the first and second derivatives of the stimulating current term.

2014]

5.1.2 Current Estimation in Real Neurons

The twin experiments on simple Hodgkin Huxley suggest that replacing the current term with an arbitrarily yet smoothly varying time dependent parameter allows for an inverse problem to estimate the current acting on a neuron based on time series voltage measurements. This procedure has the potential to allow for a passive probe into the natural currents acting on these neurons - such as from synaptic connections without the use of activity destroying voltage clamp techniques.

In order to demonstrate the validity of this method to characterizing the currents in real neurons we attempt to perform the inverse problem on the HVC interneuron data from full cell current clamp experiments in slice. [Meliza, et al., 2014] Initially a period of 1500 ms of 50 khz voltage response to a known applied current is used to estimate model parameters. These model parameters are then fixed and information about the state of the stimulating current is removed by replacing the driving current with a smoothly varying time dependent parameter. Using a window of voltage information from the same neuron[2] we estimate the stimulating current as a function of time using the modeled behavior as a filter. Estimates for the current in the same neuron data used to fit the data are shown in figure 5.2, for another time period in the same neuron in figure 5.3.

The quality of the fits to an otherwise known stimulus provides us with a good consistency check to the model. The places where the model cannot be used to predict the stimulating current, such as at around 400 ms in Fig. 5.3, may help in determining the types of features missing from the model. Despite errors in some regions, the accuracy of this fit suggests we may be able to use this method as a tool to probe unknown currents in the network in future experiments.

[2]This is the same epoch at the moment

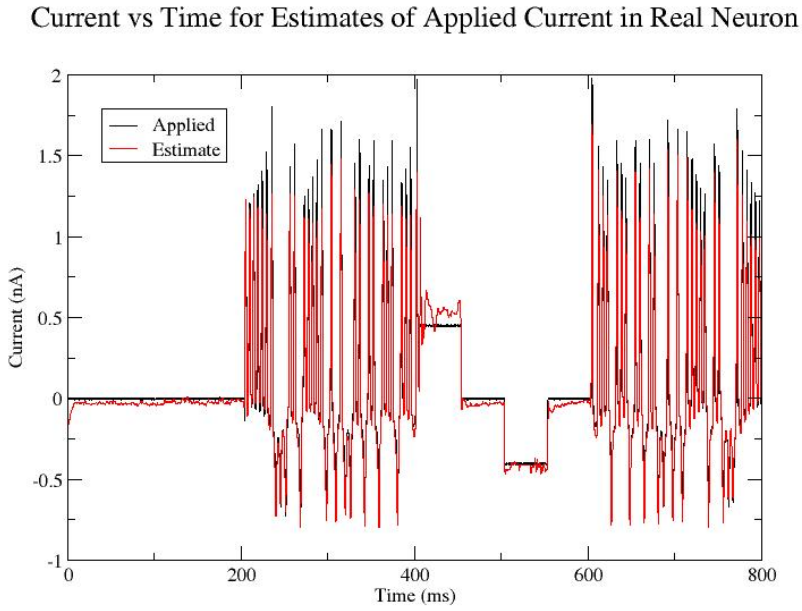

Figure 5.2: Estimates of stimulating current vs time for an identified neuron based on time series voltage measurements. The model of the neuron is fit using the measured voltage response to a known current using the variational method on the path integral. The model parameters are then fixed at the estimated values and the model is used as a filter to estimate the stimulating current used to generate the measured voltage trace. The estimated current is subjected to a continuity and smoothness constraint through a cost term to its local concavity. The voltage data epoch is the same as was used to estimate model parameters.

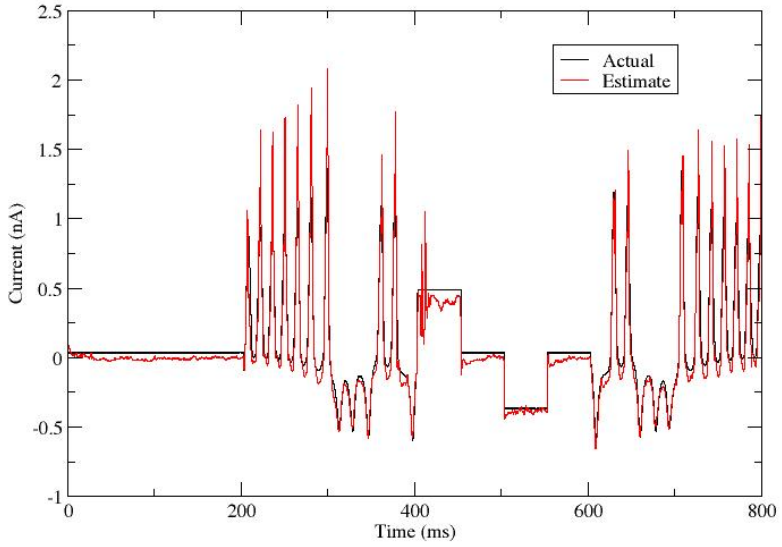

Figure 5.3: Estimates of stimulating current vs time for an identified neuron based on time series voltage measurements. The model of the neuron is fit using the measured voltage response to a known current using the variational method on the path integral. The model parameters are then fixed at the estimated values and the model is used as a filter to estimate the stimulating current used to generate the measured voltage trace. The estimated current is subjected to a continuity and smoothness constraint through a cost term to its local concavity. The neuron is the same one used to fit the model, but the data window and stimulating waveform are different.

5.2 Synapses

5.2.1 Biological Review

We have demonstrated the ability to determine the stimulus arriving in the soma of an identified neuron and now turn our attention to the sources of those currents. Neurons in a natural biological setting tend to receive stimuli from the other neurons in the network. These currents occur due to specialized linkages between pairs of cells called synapses. These connections come in two general forms: ligand gated (chemical) and electrical (gap junction). This section will provide a brief review of the general structure and function of each types of synaptic connection.

The gap junction is essentially just a pore between two cells that acts to drive both cells to a shared potential. Mathematically, these connections are simply a resistor between different regions of the cell. Despite this apparent simplicity, the twin experiments on these types of channels are non-trivial (See 5.3).

Ligand gated synapses use a small molecule as a mediator. The first neuron will release a molecule that binds to a specific type of membrane protein on the second neuron, typically from the axon terminal of one cell and the dendrites of another.[3] The binding of this molecule will open the channel to particular type of ion, inducing a current in the second cell independent of that neuron's state. This connection is unidirectional[4] as there are no ligand gated proteins on the first neuron or the release of neurotransmitters from the second. For this reason we will refer to the neuron that releases neurotransmitters as the 'pre-synaptic' neuron and the neuron that receives them as the 'post-synaptic' neuron. [Cowan, Südhof, & Stevens, 2001]

While the above is generally what happens, modeling the behavior of these connections will require more depth. A spike arriving at the axon terminal will cause an influx of Ca^{2+} ions. This increase in calcium concentration triggers the release of packets of neurotransmitters in bubbles (vesicles) contained inside the pre-synaptic neuron. These neurotransmitters then take 1 ns to diffuse across the nm sized synaptic cleft (Fig. 5.4) and bind to the corresponding membrane protein on the post-

[3] See Fig. 4.1 for where those parts are relative to the soma.

[4] There is some information that goes backwards that is related to synaptic plasticity - but the time scales over which these changes occur are long enough that we can ignore this effect at this time.

synaptic neuron. The membrane protein will then undergo a conformational change into a semi-stable open configuration allowing the flow of a particular combination of ions into or out of the cell. The ligands will stochastically fall off over time, closing the channels after which they can either bind to another cell or be reuptaken by the presyanptic neuron for later release.

There are two types of ligand gated synapses that will be considered here the AMPAergic and GABAergic synapses, each named of their respective neurotransmitter. The AMPA channel is non-selective to the types of ions it allows through and thus acts as a temporary hole between the intracellular and extracellular medium and thus $E_{AMPA} \approx 0$ mV. Because the resting state of the neuron is polarized with respect to the extracellular medium, these holes serve to depolarize the cell and are commonly referred to as excitatory synapses. The channels activated by the GABA neurotransmitter allows the flow of chloride ions into the cell and potassium out of the cell yielding a typically hyper-polarizing current - $E_{GABA} \approx -80$ mV. Because of this depolarization the synapse is typically referred to as inhibitory as the hyper-polarization suppresses firing in the post synaptic neuron. While these types of synapses are both common in a network, the synapses projecting from a given neuron (the synapses for which the neuron is the presynaptic neuron) will be either be all excitatory or inhibitory but not both. [Johnston and Wu, 1995] For this reason we will refer to cells with inhibitory/excitatory synapses as inhibitory/excitatory neurons.

5.2.2 A Model for Synaptic Currents

Similar to our work with individual neurons we develop initial stimulus and measurement routines with a simple model. We modify the Hodgkin Huxley model in equation 4.20 by adding an additional current to the post-synaptic neuron for the flow of ions governed by the synaptic activity.

$$I_s = \sum_i gE_i s_i(t)\big(V - E_{AMPA}\big) + \sum_j gI_j s_j(t)\big(V - E_{GABA}\big) \tag{5.2}$$

gE_i is the maximal conductance of the excitatory connection from the i'th excitatory (AMPAergic) cell, $s_i(t)$ is fractional activity of that synapses corresponding physically to the fraction of neurotransmitter gated ion channels in the synapse that are both

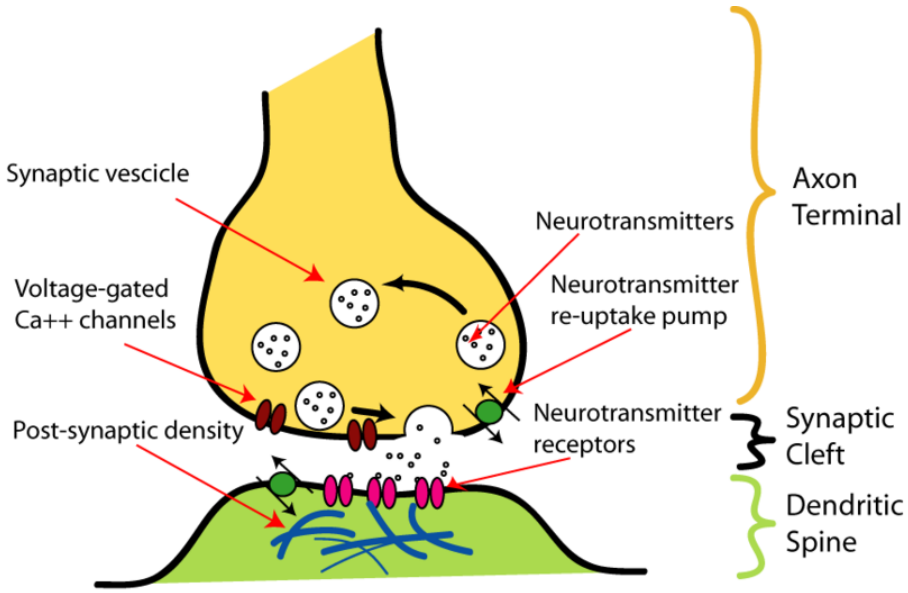

Figure 5.4: Infographic of a typical ligand gated synapse. Reproduced under GNU free documentation license.

bound with a neurotransmitter and open, and E_{AMPA} is the reversal potential for that channel - typically around 0 mV. Similarly the inhibitory synapses (GABAergic) are represented equivalently with the dynamics of s(t) having the same form and $E_{GABA} \approx -80$ mV. Since our neurons are assumed to be in isolation to other parts of the network - there will only be a single synaptic connection of type determined by the type of the pre-synaptic neuron. The dynamics of the gating particle, s(t) are assumed to behave similar in form to those of the sodium and potassium ions except that instead of being determined by the voltage of the cell they act on, the release of neurotransmitters (and thus the activation of the ligand gated channels on the post synaptic cell) are governed by the voltage of the axon terminal.

In principle we should expect a two state model with the voltage in the pre-synaptic terminal and reuptake processes governing changes in the concentration of neurotransmitter in the synaptic cleft. This concentration would then determine the expected steady state fraction of ligand gated membrane proteins on the post synaptic neuron that would be bound to a neurotransmitter and in an open configuration. This can be refined further by noting that the release of neurotransmitters is not voltage dependent, but are a function of the calcium concentration in the axon terminal -

Table 5.1: Qualitatively chosen values for the parameters for excitatory (AMPA) and inhibitory (GABA) synapses for twin experiments to develop methods for determining synaptic properties. The maximal conductance of the synapse will vary substantially between synapses. The values are chosen to qualitatively match the current responses described in the literature. [Destexhe, Mainen, & Sejnowski, 1994]

Param.	Excite.	Inhib.
E_s (mV)	0	-80
V_s (mV)	-20	-20
dV_s (mV)	5	5
C_1 (ms)	1	1
C_2 (ms)	1	3

which is a function of the local voltage gated calcium channels. While two or more state models for this behavior are more physical we have found we can qualitatively reproduce this behavior with a single variable model. This model is easier to fit - and since we are throwing away the behavior of the axon and dendrites in this model we have already shown that we are more interested in a qualitative fit than a physical one.

$$\dot{s_i}(t) = \frac{s_\infty(V_{pre}(t-\tau) - s_i(t)}{\tau(V_{pre}(t-\tau))} \tag{5.3}$$

$$s_\infty(V) = \frac{1}{2}\left[1 + \tanh\left(\frac{V - V_s}{dV_s}\right)\right] \tag{5.4}$$

$$\tau(V) = C_1 - C_2 s_\infty(V) \;\; w.\;\; C_1 > C_2 \tag{5.5}$$

5.2.3 Twin Experiments on Ligand Gated Synaptic Currents

While neurons comprise the nodes of networks of neurons - and an understanding of their dynamics is essential to characterize the behavior of a network we must also develop a similar understanding of the synaptic connections linking the nodes together. Because of their size - probing the synapses themselves is an extremely difficult proposition. In order to probe their behavior we need to develop data analysis tools to infer their properties based on their effects on the readily measurable properties of neuron behavior, namely the voltage response of the soma of a single neuron and the firing times in the soma's of the other cells in the network. This data analysis tools comes in the form of the previously demonstrated ability to estimate

the stimulating current acting on an identified neuron given a sufficiently accurate model describing its behavior and time series measurements of the voltage behavior.

It is important to note that the estimations of stimulating current are based on the currents acting on compartment the cell is being recorded from - typically the much larger soma. This is not the compartment that typically receives the actual synaptic currents - which tend to act on the dendritic compartments. The currents received by these compartments is filtered - potentially actively [Häusser et al, 2000] due to the leaky nature of the compartment and active ionic channels. Further, since the propagation of this current through the dendrite takes time, the current in the soma will be delayed relative to the initiation of the synaptic current in addition to any filtering preformed by the dendritic compartment due to diffusion time. While these features are certainly important [Golding et al. 2002] we instead ask a much simpler question, assuming we know the spike times of a neuron that projects onto our post synaptic cell, what is the typical current received by the soma? This is a relevant question because the activity in the soma determines whether an action potential will propagate into the axon compartments and stimulate other cells. [5]

The method for estimating synaptic properties relies on the use of paired recordings - simultaneous patch clamp recordings [Neher, et al., 1981] from a pair of proximally located cells. By having the ability to stimulate and record from both post and pre-synaptic neurons[6] we are able to selectively stimulate the activity of the synapse through the generation of frequently spiking waveforms in the pre-synaptic cell and magnify the effects of the stimulus on the post synaptic cell by driving the cell away from the reversal potential of the synapse. This allows for larger levels of activity in short time windows relative to the time scale of the evolution of these synaptic connections in response to our hammering due to plasticity effects. Since these experiments can potentially be preformed in slice preparations, the natural isolation will limit signals from other neurons provided the population size in the slice is below a level to allow for sustained activity. This minimizes the chance the current waveforms we record are the result of correlated activity - but limits us to

[5] Of course if it turns out that the contribution of multiple synaptic activation do not add linearly to the current acting on the measured axon, we may need to reevaluate these implicit assumptions

[6] with the possibility that they can be recurrently connected

assuming that the signals that are important for neuron response are uncorrelated at the point of the synapse, which as was noted earlier - may not be the case.

The voltage in the axon terminal is assumed to be a delayed copy of the voltage in the soma. Similarly the current arriving in the dendrites is assumed to propagate without loss to the soma of the post synaptic neuron. Without loss of generality, the time delay between dendrite and are combined into a single time delay (τ) which is typically around 5-6 ms in HVC. Obviously this is not going to be accurate given the filtering in both axon and dendrites. As we are interested in creating a predictive model for the synaptic connections based on activity in the soma, we must accept that the estimated parameter values for the kinetics of the synapse will be accounting for this filtering and may not be truly representative of the underlying physics such as the docking or undocking times of the neurotransmitters or voltage sensitivity of neurotransmitter release.

We create a twin experiment with two single compartment neurons connected by a synapse of known type. A stimulating is applied to the pre-synaptic neuron to promote frequent yet not tonic spiking behavior. This insures that any synaptic activity is varied over the data sets and allows for differentiating possible responses to different numbers of successive spikes and resting periods between bursts of activity. The post synaptic neuron is stimulated with small DC or slowly varying current to drive the neuron away from the approximate reversal potential of the synaptic current. Since the magnitude of the synaptic current scales with $(E_s - V)$ this will increase the magnitude of the signal. We generate data using a Runge-Kutta method for delay differential equations and take the voltage states for each neuron with added 0.2 mV sigma Gaussian noise additive noise taken as measurements.

While in principle connections can be bidirectional, the stimulus routine presented to the defined pre-synaptic neuron is such that it causes frequent spiking behavior. This spiking behavior limits the resolution of the driving stimulus from voltage measurements. More importantly, the stimulus in the defined post-synaptic neuron is chosen to prevent spiking activity in order to maximize our resolution of the current waveforms created by the synaptic activity. Since the activation of the soma - and by extension the axon terminal are suppressed there is little, if any recurrent activity. Even if there are recurrent activity of some form, the synaptic activity is

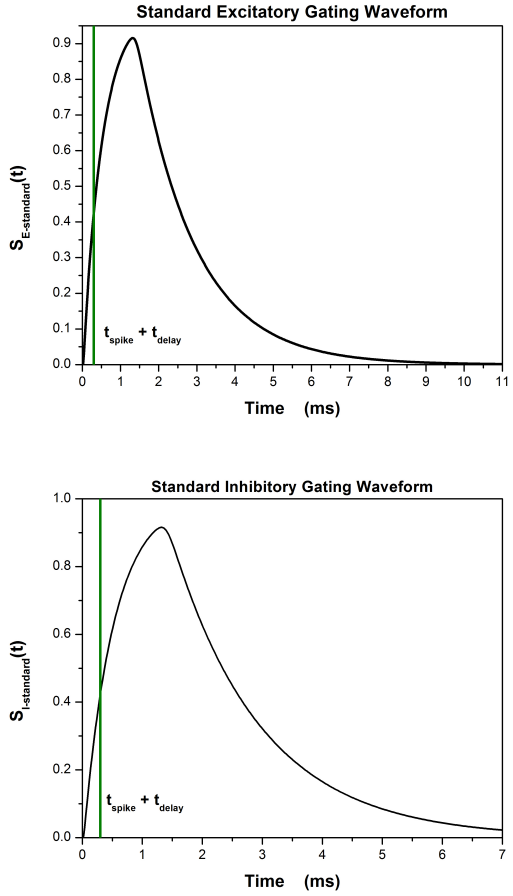

Figure 5.5: Plot of synaptic activity vs time (s(t)) for excitatory synapse (left) and inhibitory synapse (right) in response to the action potentials of the model neuron of the previous section. The green line indicates the time the peak of the pre-synaptic spike arrives at the axon terminal. Because of the physical distance between the axon terminal and soma, there is an implied time delay if the somatic voltage waveform is used as a proxy.

assumed to be a function of the pre-synaptic voltage - albeit indirectly. The actual source of the spiking behavior of in the pre-synaptic neuron is unimportant, only that we have full knowledge of the voltage response of that stimulus.

For the estimation procedure, the individual neuron parameters are set to the values determined using the single isolated neuron twin experiment in the presence of noise. Since these parameter values are functional, yet incorrect, we simulate the presence of an incorrect model and some level of dynamical noise. The two waveforms are first compared to find the time delay implicit in the model based on the firing time in the pre-synaptic neuron vs the start of voltage deviations in the post synaptic neuron. Since we are assuming that the potential for recurrent behavior is negligible or irrelevant for this experiment, by knowing the time delay in the data, we are able to shift the voltage and driving stimulus data sets to eliminate the time delay from the dynamics of the system - allowing for the model term in the cost function to remain a sum on nearest neighbor pairs of time points. Rm for each neuron is set at 1 and Rf for the regular ionic and synaptic currents are chosen to be 1e5 and 1e6 respectively. Because the actual source of the voltage behavior in the presynaptic cell is unimportant, we can represent it in the model as a stimulus driving the synaptic activity as opposed to a full many state neuron. While this is not necessary for the problem, based on the formulation of our model and on the previously demonstrated ability to determine stimulus from voltage in an identified neuron we are confident that dropping the dynamics of the presynaptic neuron does not detract from the results in any way while providing a large numerical boost. The estimates of parameter values for twin experiments on AMPA and GABA synaptic connections are shown in table 5.2. Plots of the estimates unmeasured synaptic gating variables, measured voltages and predicted behavior are shown in figures 5.6 and 5.7.

The simulated experiments for a controlled paired recording demonstrate that we can achieve accurate estimates for the behavior of both excitatory and inhibitory connections. We have also demonstrated that we can determine parameter sets that accurately represent that behavior. In some cases, such as the maximal conductance of the synapse (g_s), these parameters accurately describe the true values used to generate the data. In other cases, such as the slope of the sigmoid function (dV_s), the estimated value of the parameter is wildly different then the true value. This error

Table 5.2: Estimates of excitatory synaptic parameters in the paired recording twin experiment. Reversal potential is fixed at the true value. The inverse of dV_s is used without loss of generality.

	AMPA		GABA	
Param	Act	Est	Act	Est
g_S (mS)	0.08	0.085	0.45	0.47
E_s (mV)	0	0*	-80	-80*
V_s (mV)	-20	-23.1	-20	-19.7
dV_s^{-1} (mV^{-1})	0.2	0.08	0.2	0.138
Cs1 (ms)	1.0	1.10	1.0	1.04
Cs2 (ms)	1.0	0.91	2.0	1.82

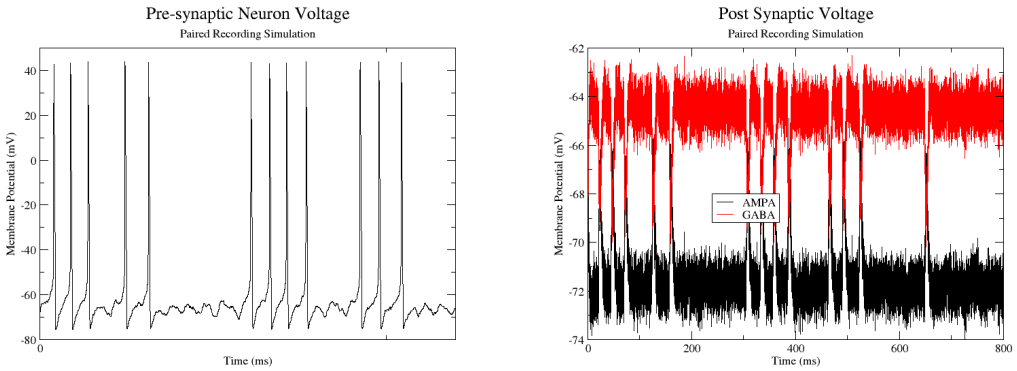

Figure 5.6: Left: Pre-synaptic voltage for both paired recordings. Right: Post-synaptic voltage for the the excitatory (red) and inhibitory (green) paired recording twin experiment. The post-synaptic voltage is modulated by an additional low frequency noise stimulus (1 nA). The pre-synaptic voltage is assumed to be the same at the axon terminal as the measured soma. The inherent delay has been temporarily eliminated due to the lack of recurrent connections. Single neuron parameters are fixed at the values estimated with noise in the previous chapter. (Tab. 4.2) A known DC current is applied to the post-synaptic neuron to drive the resting state away from the reversal potential of the respective synapse.

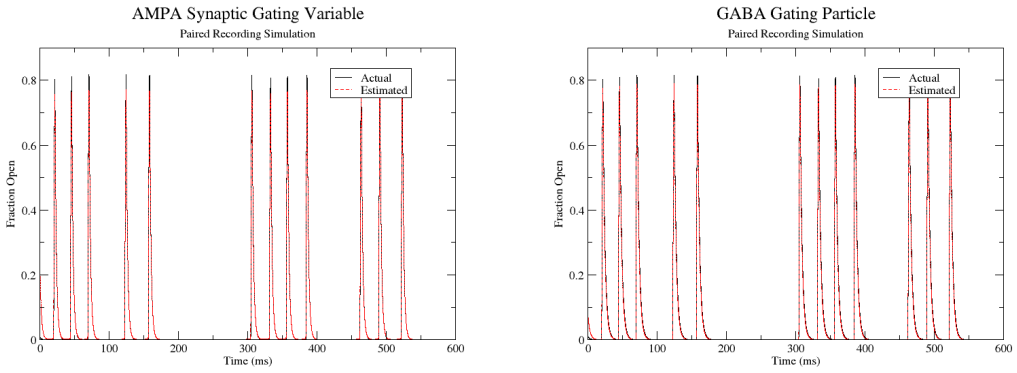

Figure 5.7: Actual (black) and estimated (red) values for the unmeasured synaptic gating variable s(t) for an excitatory synapse (left) and inhibitory synapse (right). Estimates are made based on measurements of the voltages (Fig. 5.6) of both pre and post synaptic neurons.

does not effect the trajectory of the unmeasured synaptic gating variable (Fig. 5.7) this indicates that either the model is insensitive to that parameter or the stimulus does not provide sufficient resolution in the measurements.

For the isolated single neuron twin experiment this apparent degeneracy can often be lifted by a change in the stimulus protocol to create a range of behaviors in the region of state space primarily influenced by the incorrect parameter(s). For the synaptic twin experiment, the stimuli that determine the behavior of the synapse are a direct current stimulus to the post-synaptic neuron modulating the synaptic current and an indirect stimulus through the voltage behavior of the pre-synaptic neuron modulating the synaptic activity. Unfortunately the voltage behavior at the axon terminal is difficult to control directly. The spike waveform that activates the synapse is generally independent of the stimulating current due to the orders of magnitude difference between the active sodium and potassium currents (1 μA) and applied currents (10 nA) this greatly limits the . The filtering by the axon may also smooth over any differences in spike shape - preventing even limited variation in the voltage activity at the axon terminal from the readily accessible somatic compartment.

In a real experiment we would not know in which direction(s) or type the synaptic connections would be before sticking the probes in. A work-around would be to use the stimulus routines for all possible configurations in sequence then pick

out the epoch containing the appropriate stimulus routine to preform our estimation on. This assumes that the correct synapse type(s) and direction(s) can be visually determined from the resulting voltage traces.

We have assumed throughout this that the voltage response at the axon terminal is represented by the voltage activity in the soma and that the current created by the synapse acts directly as a current on the soma. If the soma and dendrites are only preforming essentially linear filtering on these signals, the effects of this filtering may only alter some of the parameter values in the model such that while the model accurately predicts the behavior of the synapse and neurons, the model parameters do not correspond to real physical quantities. If instead the axon and dendrites provide more active, non-linear filtering, we may need to consider using multi compartment models to describe the behavior of neurons in networks.

5.3 Gap Junctions

The previous experiments dealt with the voltage activated, ligand gated synaptic connections that are common in networks. A second class of synapses, the gap junction, occurs primarily where speed is paramount, such as visual systems and reflexive escape mechanisms *(cite)*, but appears whenever cells are in close proximity. Rather than relying on the release of ligand gated proteins to activate or deactivate the adjacent cells, gap junctions are simply pores that directly link the intracellular medium of one cell to the other allowing all small molecules (like ions) to flow freely.

The pore is essentially a tube linking the two membranes together so the intracellular medium is essentially continuous between the two cells. The continuity in intracellular medium means that propagating voltage signals will travel directly through the cell membrane into the adjacent cell. Because the connection is through a smaller surface area than the average cross sectional area of the cell (Eq. 4.2) this connection is equivalent to a large resistor connecting the two 'circuits'.

$$I_{i\to j} = \frac{g_{ij}}{Area_j}(V_i - V_j) \tag{5.6}$$

The conductance of the synapse (g_{ij}) is symmetric due to conservation of charge. Since the connection is localized, the total conductance must be scaled by the sur-

face area of the cell before being applied like the per unit area conductances of the Hodgkin-Huxley model. As the compartments are linked, the reversal potential of the channel is simply the current voltage of the linked compartment.

For the ligand gated synapses, the activity of the synapse was driven by the voltage state of the pre-synaptic axon terminal. If there are recurrent connections they could potentially alter the pre-synaptic voltage waveform altering the activity. We ignored this possibility in the earlier twin experiments because in all cases the time scale of the communication (5 ms in each direction of simple travel time in the axon) is much longer than the voltage response of the synapses. More importantly, by measuring the pre-synaptic voltage we learn the effects of the stimulus driving that neuron even if we do not know the full source of the currents driving it. For the gap junction the connection speed is generally[7] not mediated by any intermediate gating particle. This leads to very fast, bidirectional coupling between both compartments that negate our ability to ignore the full dynamics of one of the neurons.

The astute reader will have noticed we have used the words neuron and compartment interchangeably in this section. This is not in error. While gap junctions do connect pairs of neurons, the mathematical structure can be used to link different components of neuron together to expand on the single compartment model with the coupling now representing the resistance of a length of intracellular medium. The single neuron models used in the past assume that the soma of the neuron is representative of the full state of the neuron modulo possible time delays due to the finite propagation speed of signals throughout the cell.

Obviously this can't be completely correct - as the neuron contains dendrites and an axon that project onto and from the soma respectively. While signals generally go from end of the neuron to the other, these additional 'compartments' can and do have different types of channels and channel densities than present in the soma. Since these compartments are directly coupled - the feedback between the compartments may need to be considered to determine the full behavior of a real neuron. For the single neuron twin experiments and experiments on real neurons, the

[7] In some cases the channel can be closed in response to certain activity patterns - such as the death of one cell.

current is applied directly to the measured component. For networked neurons, the connections - and thus applied currents, are applied via an axon compartment in the pre-synaptic neuron to a dendritic compartment of the post synaptic neuron. Given that these compartments are not the measured somatic compartment, understanding how these compartments interact is going to be essential for characterizing the behavior of networked neurons.

We approach the twin experiment similar to the ligand gated synapses we assume that the two neurons have known properties which is represented by setting the Hodgkin-Huxley parameters to those estimated in the noisy isolated neuron twin experiment. Since the connection is fast and bidirectional, we use the full Hodgkin Huxley model for both compartments. While the types of currents can be in principle be different, for example, between dendritic and somatic compartments we choose to assume that the types and densities of channels in both compartments to be the same. While the channels may be the same, we do not assume that the area of the two cells are the same. The surface area of different compartments may be vastly different - so we leave that parameter free for both neurons. To stimulate activity, we apply the same stimulus used for the isolated single neuron (Fig. 4.5) to one of the compartments. Data was generated using stochastic RK with added low frequency noise to simulate background noise in the system. Additional 0.5 mV variance guassian noise was added to the extracted voltage states to simulate noise in measurement apparatus.

With measurements of the voltage in both compartment we can estimate the connectivity of the resistive synapse or length of intracellular medium along with all unmeasured states. While there are errors in the estimated values of the area and connectivity parameters (Tab. 5.3), the values are sufficiently accurate to allow for predictions of future states (Fig. 5.8)

While the additional resolution may be possible for paired recordings of two neurons, there are several technical issues that may make even this procedure as applied to multiple compartments of a single neuron. Currently, patch clamp recordings on dendritic compartments are possible [Davie, et. al. 2001]. Additional difficulty is provided by the requirement that a second recording be made from the soma of the same cell. This recordings have been done in large Purkinje cells [Roth, 2001],

Table 5.3: Parameter estimates for the gap junction or two compartment neuron twin experiment. The other Hodgkin Huxley neuron parameters are fixed at the estimated values of the noisy single neuron twin experiment for both cells. Estimates are conditioned first on voltage measurements of both compartments then on only one compartment.

Param.	Actual	Both Meas.	One Meas.
$Area_1^{-1}$ (cm^{-2})	0.8	0.81	0.864
$Area_2^{-1}$ (cm^{-2})	2.0	1.96	0.667
g_{Gap} (mS)	5e-2	5.1e-2	0.114

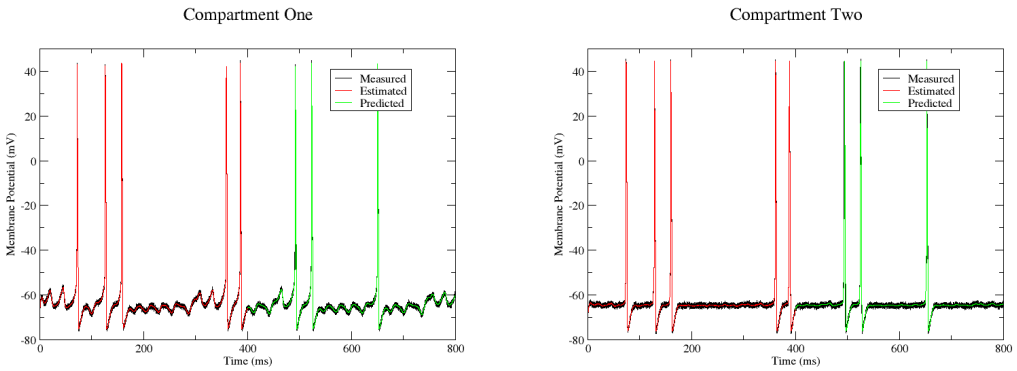

Figure 5.8: Plot of voltage vs time for gap junction and two compartment twin experiment with both compartments measured. The Hodgkin Huxley neuron parameters are fixed at the estimated values of the noisy single neuron twin experiment for both cells. Predictions are generated by integrating the model forward using the estimated parameters and final state of the system as initial conditions. Note that while all spikes induced in the excited compartment lead to spikes in the other compartment, the sub-threshold behavior is filtered out.

but have not been preformed in the smaller neurons of the HVC system. Given these types of recordings are both difficult and invasive there is no guarantee that multiple measurements of each neuron will be possible for a given system let alone one. The resistively coupled system transmits information from the stimulated to the coupled neuron in the form of current proportional to the voltage difference. The coupled neuron than returns a current in response that is again proportional to the voltage difference between the two compartments. Since we have demonstrated earlier (Sec. 5.1) the ability to estimate the external currents acting on an identified single compartment neuron from voltage measurements, those measurements should provide voltage information about the unmeasured compartment, which will turn provide information about the unmeasured gating variables and possibly a few parameters.

As a test of this ability to predict unmeasured states we consider the same system except we only use measurements of the stimulated compartment (compartment one) to estimate states and area and connectivity parameters. The data sets are identical to the paired recording simulation. The estimates of the parameters (Tab. 5.3) show that the area of the of the measured compartment can be estimated accurately but not the area of the unmeasured compartment and resistive conductance. The values of the resistive conductance and area are such that while the voltage in the other compartment is poorly estimated (Fig. 5.9) the current acting on the measured compartment from the unmeasured compartment is very similar. This is potentially reflected in the accuracy of the predictions of the measured component. Alternatively, the effects of the feedback may be such that this return effect is insignificant when compared with other currents in the cell. A different set of compartment sizes and coupling strength may have a larger feedback - resulting in more accurate estimates, however we cannot rely on the parameters being in useful regions for our estimation procedure to work.

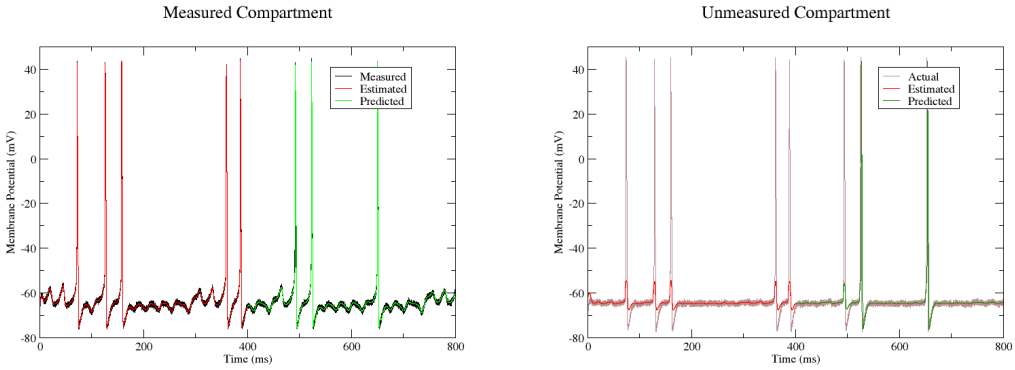

Figure 5.9: Plot of voltage vs time for gap junction and two compartment twin experiment with one compartment measured. The Hodgkin Huxley neuron parameters are fixed at the estimated values of the noisy single neuron twin experiment for both cells. Predictions are generated by integrating the model forward using the estimated parameters and final state of the system as initial conditions. While the measured compartment's future behavior is well predicted, no spikes are estimated in the second compartment, even though the prediction indicates it should be spiking.

Chapter 6

Neural Networks

Cognitive function in both humans and animals are a complex dance amongst many interconnected subnetworks, each preforming some function then passing on the output - in the form of a series of spikes, to other subnetworks of the brain. Understanding how these subnetworks interact requires an understanding of how each particular subnetwork functions. The function of each subnetwork is determined by the types of cells present in the network and the connections between them. We have demonstrated in previous sections the ability to characterize the behavior of individual cells and the kinetics of individual synapses using voltage recordings from single or pairs of cells. These estimates allow us to constrain our estimation of network behavior by fixing the behavior of the network nodes and connections kinetics to known values, leaving the only the architecture of the connections unknown. This chapter will discuss the use of the path integral method as a means to determine the connectivity of a small network of neurons in a twin experiment conditioned on the types of data that would be available using current experimental techniques.

6.1 State estimation through spike timing

The simulations preformed in the two previous chapters were made using time series recordings of the membrane potential of either one or two cells. While there are several means for obtaining these values such as sharp electrode or patch clamp techniques, these methods tend to use probes that are large compared to the size of

the measured neuron. For determining the behavior of individual or pair of cells, the size of the probes and how they fit into the network tissue can generally be ignored. As the number of measured components in system increases, there is decreasing space for additional probes. For networks of neurons we need to consider alternative measurements to determine the state of the neurons that are less intrusive to provide greater breadth at the expense of depth.

The ability to use a potential reduced measure of the state of the system depends on the synchronization properties of the measured system. We have previously determined that the voltage information is sufficient to estimate the full state and parameters of an isolated neuron. Given that a large component of the information in the measurements is used to estimate the parameters fixing the parameter values near their true values may reduce the frequency at which measurements are required to estimate the state of the system.

To determine how frequent we must make measurements, we return to the structure of the Hodgkin Huxley model. For hyper-polarizing and weakly polarizing stimulus, the Hodgkin Huxley model evolves to a stable fixed point at the resting state of the neuron. For larger static depolarizing currents, the neuron will fire tonicly, (limit cycle) - with frequency increasing with stronger stimuli. At very large stimuli, the sigmoid functions saturate at their high voltage state (open for m, n, closed for h) and system returns to a stable fixed point. Since this requires the stimulus to overwhelm the potassium current this particular transition can be ignored as that strong a stimuli will kill the cell.

Given that the state of the system is always going to be on some trajectory with a stable fixed point or limit cycle, one might naively assume that this means the full state of the system can be estimated at all time points without measuring any states of the system. Using the parameter values estimated from the noisy single neuron twin experiment we estimate the unmeasured (all) states of the system using the variational method on the path integral, the results of which are shown in figure 6.1. While the fixed point transitions in the subthreshold region are accurately predicted, the variational method fails estimating transitions from fixed point to limit cycle.

The failure to estimate the now unmeasured voltage during these transitions can be attributed to landscape of the cost function. As was derived earlier (Eq. 3.12)

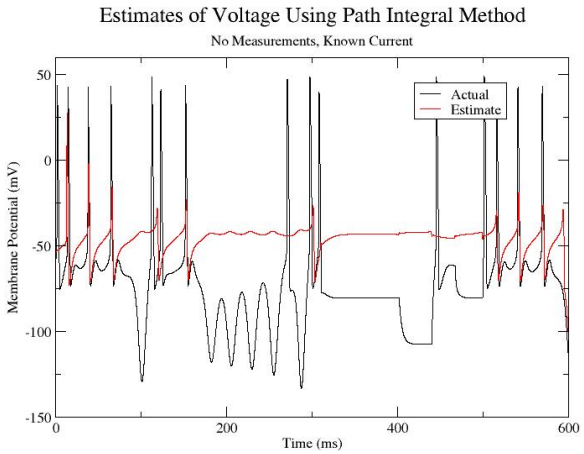

Figure 6.1: Estimates of voltage vs time for a simulated isolated neuron with known stimulus but no measurements. The instabilities at the spiking threshold create a local minimum that prevents the path from evolving from subthreshold to spiking behavior. Despite the lack of measurements the subthreshold behavior is predicted extremely accurately, indicating that measurements are not needed in these regions.

the variational step:

$$\frac{dx_i(t)}{ds} \approx \sum_{j=1}^{D} \left[\delta_{ij} + dtJ_{ij}(\mathbf{x}(t))\right] Rf_j(t) \tag{6.1}$$

is stable with respect to perturbations only in regions where the system is not *locally* chaotic. As can be seen in a colored coded phase plot of the m and h gating variables for a series of pulses of current with steadily increasing magnitude (Fig. 6.2) the transition from subthreshold to spiking behavior and back is locally unstable - which as per Eq. 6.1 corresponds to unstable regions of the cost function landscape.

The instabilities, as defined by areas with a local positive Lyapunov exponent, occur only at the Hopf bifurcation at the spiking threshold. This indicates that in order to estimate the full state of a neuron at all times using the path integral method, we will need some measurements around the spikes to drive the state of the system away from the local minima created by that bifurcation.

When a neuron fires, the action potential creates sufficient activity in the extracellular medium to be detected by probes adjacent to the neuron. Because they do not need to apply stimulus nor maintain a seal to the cell membrane, these extracellular probes can be much smaller than the probes used for full cell recordings. Rather

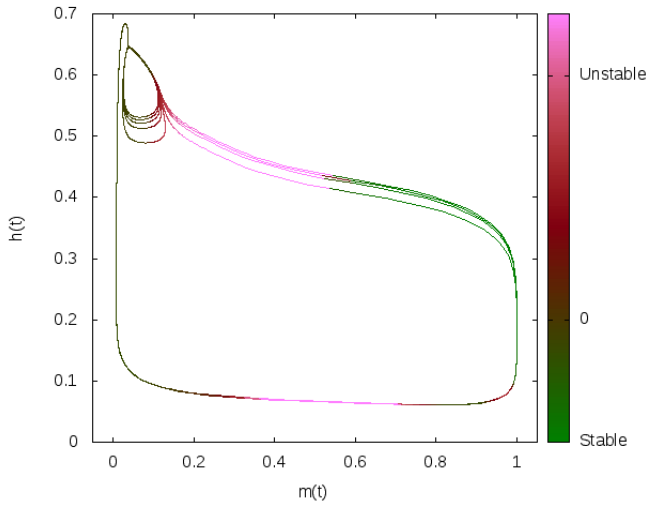

Figure 6.2: Phase plot of the sodium inactivation vs activation variable colored according to the local Lyapunov exponent. The system is stimulated by a series of pulses of increasing strength until a spike is caused. Between spikes the system sits at the fixed point in the upper left corner. The system diverges when sufficient current is applied due to a Hopf bifurcation. Note that the repolarization phase also passes through an unstable region.

than recording from one or two cells, as is the case for the full cell recording techniques, these extracellular measurements can be made simultaneously for hundreds of cells. [Buzsaki, 2004] See appendix B for an in depth treatment of the physics behind these types of measurements.

Extracellular recording techniques provide us with the ability to determine the firing times of a large number of neurons in small network. Since the spike timing information is not (directly) about the membrane potential we must modify the firing times into a time series voltage waveform that correspond to that spike. The stimulus currents are of order magnitude 1-10 nA while the sodium and potassium currents are of order 1 μA. This means the general shape of a spike once the sodium channel opens is generally independent of the underlying stimulus. This allows us to create a stereotype spike waveform by average multitude of spikes generated by the neuron - as shown in Fig. 6.3.

Note that in real neurons there is some variation in spike shape due in part to the statistical nature of the finite number of gating proteins [Anderson et al, 2014] and variations in slower driving currents that are not present in the simple model. This

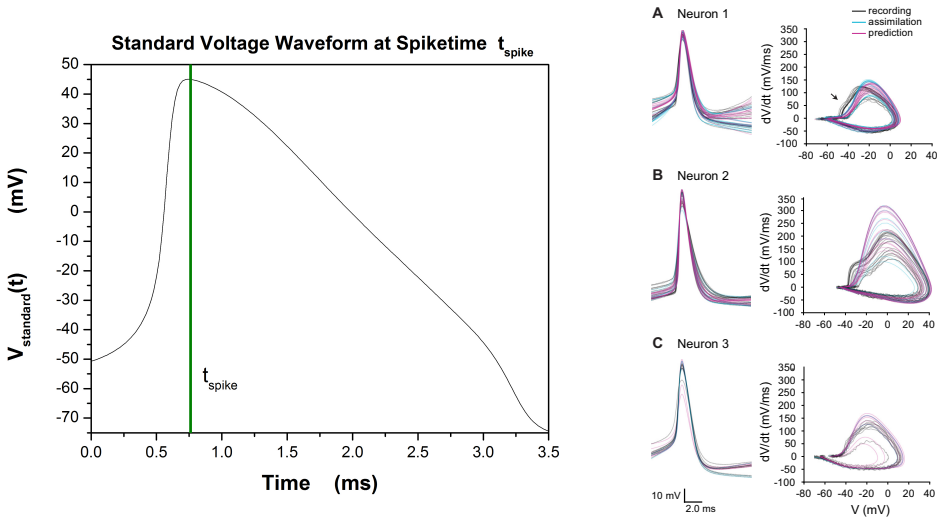

Figure 6.3: Left: The voltage activity in a neuron follows a stereotyped waveform that is generally independent in shape to the stimulus. The voltage trajectory of many spikes is averaged together to create this stereotyped spike. The measured spike time is roughly at the vertical line. Right: Spike waveforms from HVC interneuron. There is substantial variation in spike shape in real neurons, this variation can potentially be represented by a relatively weaker coupling to the average waveform shape. (Reproduced by permission from [Meliza, et al., 2014])

may necessitate relatively weaker coupling to the resulting spike waveforms (smaller Rm) and/or context dependent spike shapes. This is an area for potential future work.

The stereotyped spikes can be combined together to create a discontinuous voltage waveform that provides reasonably accurate voltage information during the action potentials. This waveform is implemented by making R_m a function of time, with $R_m = 0$ outside the stereotyped spiking waveforms. We repeat the previous twin experiment with the addition of the measured spike times. The estimates of the voltage behavior with the addition of this information is shown in Fig. 6.4. The true behavior is in black, the estimated behavior is in red, and the stereotyped spike waveforms are in green. There are no measurements outside the green region.

Note that as defined, this twin experiment does not correspond to a real experiment. The ability to apply a known stimulus waveform comes from the full cell patch clamp. This exercise simply demonstrated that the addition of spike timing information to a neuron in which model parameters and stimuli are known is sufficient

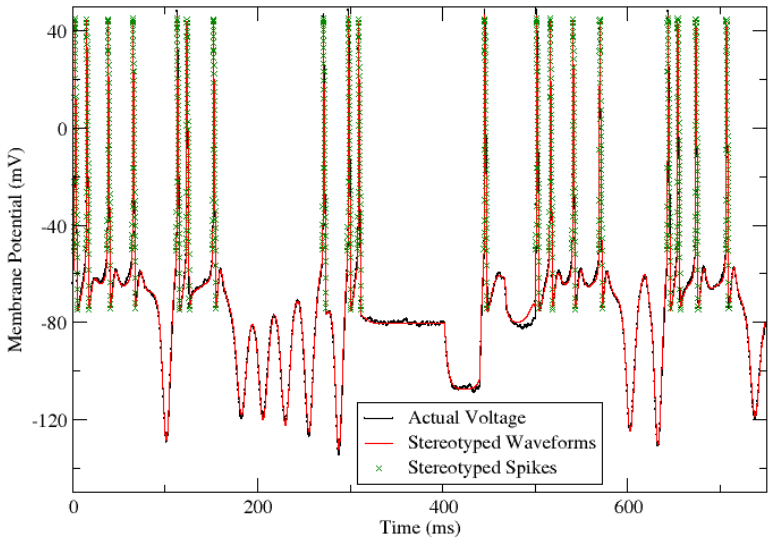

Figure 6.4: Plot of voltage vs time for an isolated neuron with known stimulus with measured spike times. Spike times are used to position a series of stereotyped voltage waveforms (green). These simulated measurements are close enough to the real activity to overcome the instability around the spiking threshold, allowing the path integral method to track the system at all times.

to estimate the full state of the system.

Given known stimulus and known model there is generally no reason to estimate the unmeasured state of the system. The global regularity of the Hodgkin Huxley model (and by implication, the neurons described by that model) means that for any physically realistic initial condition, the true system and the model system will synchronize absent any measurements coupling them together. However, if we do not have information about some parameters in the system, the feed forward estimation is less effective. As was discussed previously (see 3.4) feed forward methods like Kalman filters have difficulties with parameter estimation because they can only use local information. The path integral method allows us to leverage the entire data set for parameter estimation, but this requires an initial demonstration that the states can be estimated with fixed parameters.

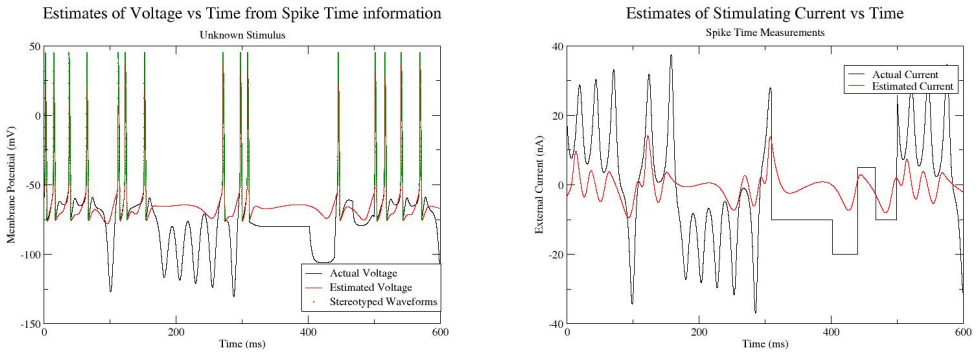

Figure 6.5: Estimates of voltage vs time (left) and current vs time (right) for an identified neuron with unknown stimulus conditioned on stereotyped voltage waveforms positioned at the spike times. Absent any voltage or current information, the voltage and current will bias towards a resting steady state. The presence of spikes in the data is a good indication of depolarizing stimuli to the neuron just prior to the spike, which is reflected by positive estimated currents in these regions. The wavy nature of the stimulus estimate - while appearing to match some general shape to the sub-threshold currents, is more a function of the smoothing algorithm that puts a cost on the second time derivative of the current than an ability to track hyper-polarizing currents. Note that the pattern is non-trivial in that the closely spaced pair of spikes is correctly estimated by a single depolarizing current waveform.

6.2 Standardized Currents

While spike timing information provides us with sufficient information to estimate the full voltage behavior of an identified neuron with full knowledge of the stimulus, it does not provide sufficient information to simultaneously estimate the neuron state and stimulus. The estimates of voltage and current for a neuron with known spike times and unknown driving stimulus are shown in figure 6.5. The stimulus to these neurons comes from the synaptic connections between the many cells. As we have little or no prior information about if or how strongly particular neurons are connected we have little information about the stimulus into each neuron. This means we lack the full voltage information needed to estimate the stimulus and lack the stimulus information to estimate the full behavior from the spike times. Provided that spike times alone are insufficient information, additional information or constraints are needed to allow for a measure of spike times must be used to estimate the connection architecture of a network of neurons.

The constraint we make is that we will assume that the stimulus each neuron receives consists of the weighted sum of all the synaptic currents caused by spikes in all the other neurons in the network. If we can make the assumption that the shape (but not strength) of the synaptic current is determined solely based on the type of synapse and the pattern of spiking in the pre-synaptic neuron that created it, the current acting on each neuron is now a linear combination of synaptic activity waveforms in the synapses:

$$I_s(t) = \sum_i g_i s_i(t)(V - E_i) \tag{6.2}$$

While the strength of the synapse from neuron i (g_i) is unknown, the activity of the synapse is a function of the (measured) spiking activity in the pre-synaptic neuron - see figure 6.6. As was demonstrated in our twin experiments with paired recordings (see 5.2.3), the activity waveform in response to a spike (or pattern of spikes) in the pre-synaptic neuron - $s(t - t_{spike})$ would be known. This allows for the activity waveforms s(t) to be constructed in pieces from the combination of the spike times and the approximate time delay implicit in modeling physically extensive objects with point models. This simplifies the search from a N*T dimensional search - where N is the number of synapses and T the number of times, into an N dimensional search over the linear combination of weights. We examine a pair of simple networks consisting of a handful of neurons to test whether this assumption provides sufficient constraint.

6.3 Two Neuron Examples

We begin with a network of two neurons with mutual excitatory connections (Fig. 6.7). We stimulate Neuron 1 with a known current and observe the timing of spikes in each neuron over a 600 ms window. The simulated data sets for the neuron voltage and synaptic activity based on these spike times and the generalized waveforms for each are used to estimate the maximal conductances of the synaptic connections in the network along with all dynamical variables of the network at all time points. The estimates and predictions of voltage versus time for each of the two neurons are shown in Figure 6.8 and parameter estimates are shown in Table 3. The accuracy of these estimates is about 10% or better.

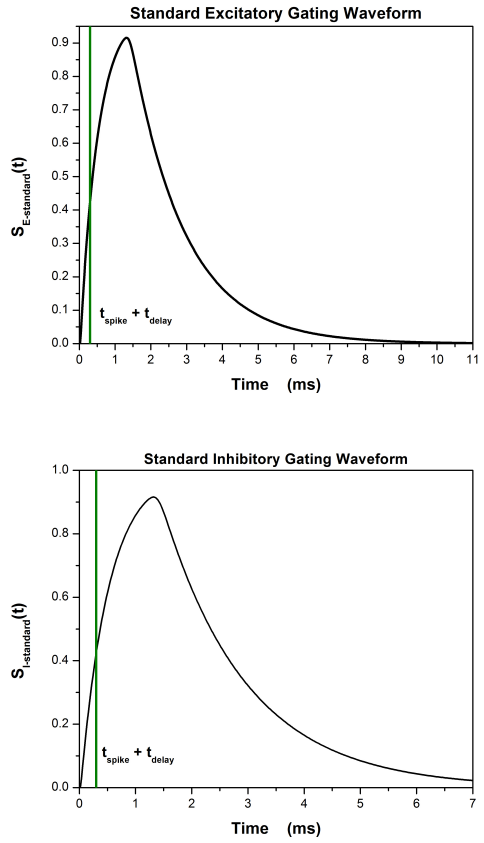

Figure 6.6: Standardized gating-variable waveforms for excitatory **Top Panel** and inhibitory **Bottom Panel** synapses. For each observed spike time the appropriate waveform is added as a measurement of the associated synaptic gating variable $S_I(t)$, with the time of the spike aligned to the vertical green line. each observed spike time a standard waveform for the synaptic gating variable $S_I(t)$ just before and just after the spike time is used as part of our data set.

Table 6.1: Estimates of the maximal conductances of the two excitatory synapses in the two neuron network; Fig. 6.7

Parameter	Estimated Value	Actual Value
$g_{E1\to 2}$ (mS/cm^2)	0.280	0.23
$g_{E2\to 1}$ (mS/cm^2)	0.197	0.18

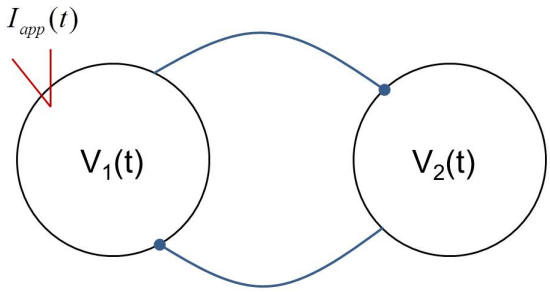

Figure 6.7: Two HH NaKL neurons mutually connected by excitatory synapses. A current $I_{app}(t)$ is injected into Neuron 1 and the spike times of both neurons are recorded. Inhibitory connections are denoted by open circles, excitatory connections by closed circles.

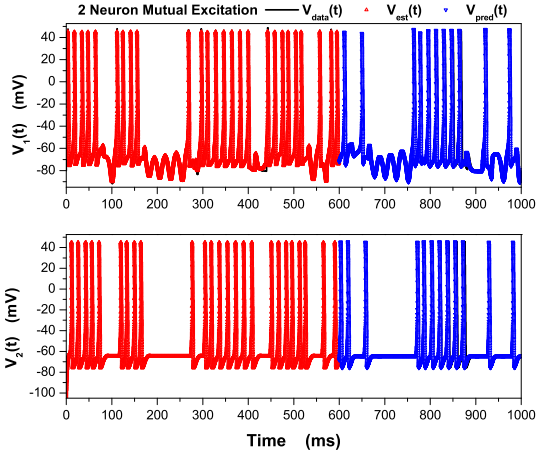

Figure 6.8: Voltage estimates and predictions for a two-neuron network with mutual excitation between the neurons. Current is injected into Neuron 1 and the spike times of each neuron are recorded. In data assimilation, for each spike time a standardized spike waveform (Fig. **??**) is assigned to the presynaptic cell, and a standardized synaptic gating-variable waveform (Fig. 6.6) is assigned to the postsynaptic cell. All parameters are fixed in the model except the maximal conductances of the excitatory connections from neuron 1 $\rightarrow$ 2 ($g_{E\,1\rightarrow 2}$) and from neuron 2 $\rightarrow$ 1 ($g_{E\,2\rightarrow 1}$). The known simulated voltages are shown as black lines, the estimated voltages by red up triangles, and the predicted voltages as blue down triangles. Neuron 1 is shown in the **Top Panel** and Neuron 2 in the **Bottom Panel**.

Table 6.2: Estimates of the maximal conductances of the excitatory synapse and the inhibitory synapse in a two neuron network (Fig 6.9).

Parameter	Estimated Value	Actual Value
$g_{E\,1\to2}$ (mS/cm^2)	0.156	0.15
$g_{I\,2\to1}$ (mS/cm^2)	0.766	0.5

While the ability to estimate excitatory currents is essential, much of the behavior of networks of neurons is regulated by inhibitory connections. While excitatory connections lead to activity that can be measured, inhibitory connections express themselves through the absence or reduction in activity. While we do not have the ability to probe sub-threshold behavior with spike timing measurements, we do know from the lack of spiking activity quite a bit about the approximate state of the system - that it is below the spiking threshold. As the neuron is not spiking, the sodium activation particle - m(t) will tend to be small until sufficient external current drives the neuron into a sufficiently depolarized state to activate it. By adding a constraint to the m gating particle outside of spikes, we create an effective upper bound on the voltage behavior. Given the close relationship the voltage has to the current, this constraint allows us to essentially say the stimulating current must be less than some value at this time for their to be a non-spike. Given that the locations of spikes will provide some information about the minimum strength of the excitatory currents - this in turn gives some information about stimulus outside regions of spiking. The lack of activity in these regions provide some information about the needed inhibitory strength.

Again, we start with the most simple case: a two-neuron network with one inhibitory synapse and one excitatory synapse (Fig. 6.9). We stimulate Neuron 1 with a known current, observe the timing of spikes in each neuron over a 600 ms window, and estimate the maximal conductances and dynamical variables as before (Fig. 6.10). Note that the firing rate in both neurons is lower than in Figure 6.8 due to inhibition by Neuron 2. The estimates of synaptic conductances are shown in Table 4. The estimate for the inhibitory synapse is less accurate than for the excitatory synapse, perhaps because the inhibition does not affect spike rate or timing when Neuron 1 is inactive.

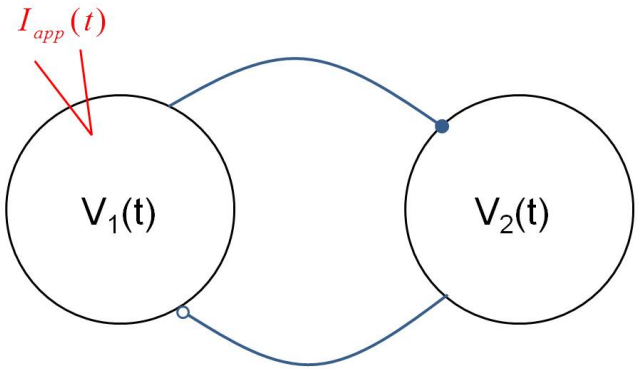

Figure 6.9: Two HH NaKL neurons mutually connected by ligand-gated synapses. The connection from Neuron 1 to Neuron 2 is excitatory (lines with full circles); the connection between Neuron 2 and Neuron 1 is inhibitory (lines with open circles). A current $I_{app}(t)$ is injected into Neuron 1 and the spike times of both neurons are recorded.

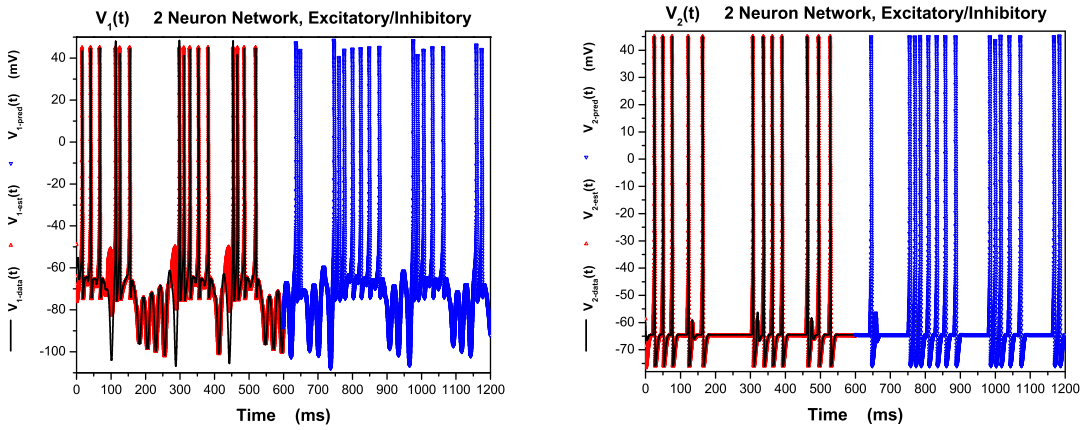

Figure 6.10: Voltage estimates and predictions for a two-neuron network with excitatory and inhibitory synapses. Current is injected into Neuron 1 and the spike times of both neurons are used to estimate synaptic conductances and dynamical variables. The known simulated voltages are shown as black lines, the estimated voltages by red triangles, and the predicted voltages as blue triangles. Neuron 1 is shown in the **Top Panel** and Neuron 2 in the **Bottom Panel**.

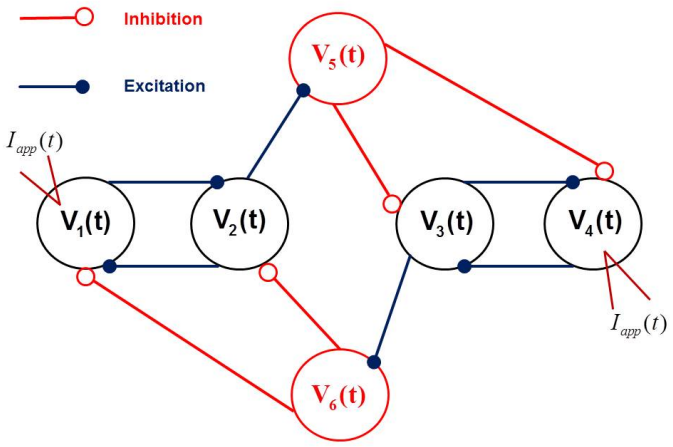

Figure 6.11: Network diagram for the six-neuron network used for estimating network connectivity. Neurons 1–2 and 3–4 form two oscillators that each activate one of the inhibitory cells (5 and 6) to suppress the activity of the other oscillator.

6.4 Estimating Connection Strengths in a Known Network

In general, neural networks of interest contain many neurons connected by many synapses, typically of order thousands or more. Networks of this size are too large to be considered using our current computing resources so we start with a smaller network consisting of six neurons connected by 10 synapses to demonstrate the methods. The model network consists of four excitatory neurons arranged in two pairs of mutually exciting cells. These cells in turn activate one of the inhibitory cells, which suppresses the other pair of cells. This configuration yields a pair of mutually inhibiting oscillators (Fig. 6.11). Known currents with low-frequency polarizing and depolarizing content are injected into neurons 1 and 4 (Fig. 9) to ensure that both states are explored. While this would not be possible with standard extracellular recording techniques we are limited to systems that have activity to measure, and that requires stimulation.

As with the simpler networks, data are generated by integrating the neuron model forward with known driving currents waveforms to stimulate activity. The spike times are extracted from the simulated voltage data to generate datasets for

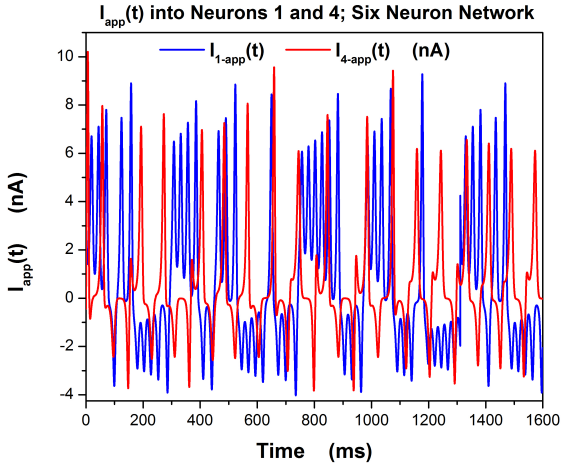

Figure 6.12: Driving currents for six neuron network. The current in neuron one is in blue and the current applied to neuron 4 is in red.

assimilation with standardized spike voltage and synaptic gating-variable waveforms. States and parameters are estimated from 1.6 seconds of data (about 50 spikes per neuron), setting $R_m = 10$ for the windows around each simulated action potential and 0 elsewhere, and $R_f = 10$ for V and $R_f = 10^5$ for all gating variables. At the end of the assimilation window, the estimated states and parameters are used to predict the response of the network to further stimulation. Figure 6.13 shows the estimated and predicted voltages for two of the neurons in the network. The estimated voltages closely match the simulated values, as do the predictions, although the predictions diverge from the data around 2100 ms. This divergence reflects the bistable nature of the network, but eventually the driving currents cause the models to re-synchronize.

Estimates for the synaptic conductances (Table 5) are within 10% of the correct values for excitatory synapses and within 25% of the values for the inhibitory synapses. The difference in relative error between synapse types suggest that spike time data constrain the excitatory coupling strengths more than the inhibitory strengths, at least with the injected current forcing we have selected.

Table 6.3: Synaptic maximal-conductance estimates for six-neuron twin experiment with fixed architecture. Only the maximal conductances among the neurons are estimated. All conductances have units mS/cm^2.

Parameter	Estimated Value	Actual Value
$g_{E\,1\to2}$ (mS/cm^2)	0.163	0.15
$g_{E\,2\to1}$ (mS/cm^2)	0.152	0.15
$g_{E\,2\to5}$ (mS/cm^2)	0.168	0.15
$g_{E\,3\to4}$ (mS/cm^2)	0.167	0.15
$g_{E\,3\to6}$ (mS/cm^2)	0.168	0.15
$g_{E\,4\to3}$ (mS/cm^2)	0.164	0.15
$g_{I\,5\to3}$ (mS/cm^2)	0.786	0.7
$g_{I\,5\to4}$ (mS/cm^2)	1.11	0.8
$g_{I6\to1}$ (mS/cm^2)	0.968	0.7
$g_{I\,6\to2}$ (mS/cm^2)	1.04	0.8

Table 6.4: Synaptic maximal-conductance estimates for six-neuron twin experiment with unknown architecture. The maximal conductances for all possible synapses are estimated, and connectivity is pruned by assuming maximal conductances less than 0.001 mS/cm^2 indicate no direct connection is present. All conductances have units mS/cm^2.

Param.	Est.	Act.	Param.	Est.	Act.
$g_{E\,1\to2}$	0.164	0.15	$g_{E\,2\to1}$	0.154	0.15
$g_{E\,1\to3}$	0	0	$g_{E\,2\to3}$	0	0
$g_{E\,1\to4}$	0.015	0	$g_{E\,2\to4}$	0	0
$g_{E\,1\to5}$	0.022	0	$g_{E\,2\to5}$	0.175	0.15
$g_{E\,1\to6}$	0	0	$g_{E\,2\to6}$	0	0
$g_{E\,3\to1}$	0	0	$g_{E\,4\to1}$	0	0
$g_{E\,3\to2}$	0	0	$g_{E\,4\to2}$	0.002	0
$g_{E\,3\to4}$	0.166	0.15	$g_{E\,4\to3}$	0.164	0.15
$g_{E\,3\to5}$	0	0	$g_{E\,4\to5}$	0.006	0
$g_{E\,3\to6}$	0.175	0.15	$g_{E\,4\to6}$	0.036	0
$g_{I\,5\to1}$	0.030	0	$g_{I\,6\to1}$	1.058	0.7
$g_{I\,5\to2}$	0	0	$g_{I\,6\to2}$	1.215	0.8
$g_{I\,5\to3}$	0.754	0.7	$g_{I\,6\to3}$	0	0
$g_{I\,5\to4}$	1.288	0.8	$g_{I\,6\to4}$	0.175	0.15
$g_{I\,5\to6}$	0	0	$g_{I\,6\to5}$	0.275	0

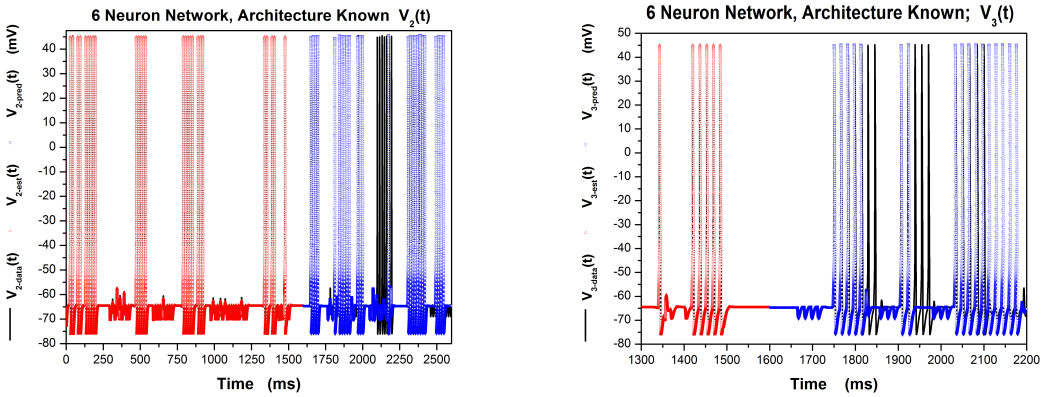

Figure 6.13: Voltage estimates and predictions for Neuron 2 **Top Panel** and Neuron 3 **Bottom Panel** in the six-neuron network twin experiment with known network architecture. In both plots, the simulated voltage is shown in black, the estimated voltage based on spike times is in red, and the predictions based on the final states and parameters of the estimates are in blue.

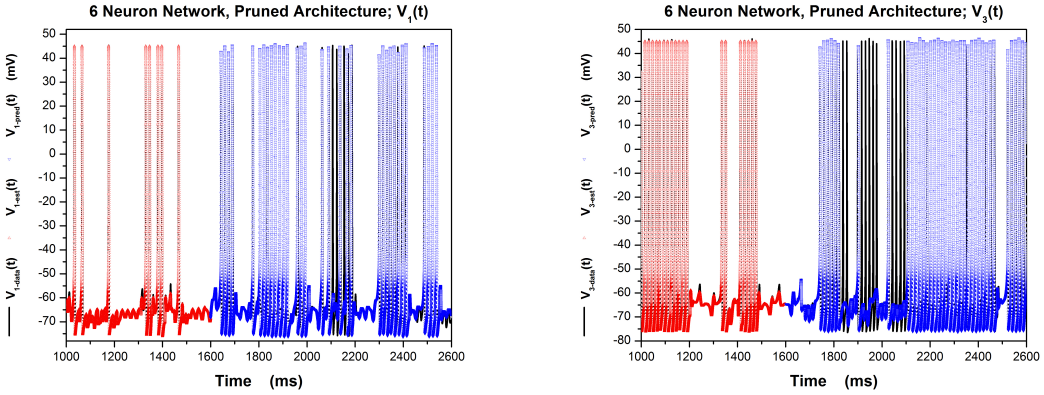

Figure 6.14: Voltage estimates and predictions for Neuron 1 **Top Panel** and Neuron 4 **Bottom Panel** in the six-neuron network twin experiment with unknown network architecture. The all-to-all initial network is pruned by the estimation procedure. The simulated voltage is shown in black, the estimated voltage during the assimilation is in red, and the predicted voltage based on the final states and parameters is shown in blue.

6.5 Pruning an All-to-All Network

In most cases the connections between neurons in a biological network are not all known when we begin our analysis. Here we ask whether it is possible to determine a network's functional architecture by measuring the spike times of neurons in a network stimulated with a known current. This means that any estimations of network connectivity would need to take into account all possible synaptic connections. In principle, this should not present any additional problems as a non-existent connection is equivalent to a connection with a zero maximal conductance.

Using the same six-neuron network as in the previous section, the estimation procedure is modified to allow all-to-all synaptic connections as an initial guess. It is assumed that we have some knowledge of cell type, namely whether a neuron exhibits inhibitory or excitatory behavior, but no information about the physical connections that are present. Figure 6.14 shows estimated and predicted voltage for one of the neurons in the network, demonstrating that the spike times are sufficient to constrain the estimated model so that it produces nearly identical patterns of activity as the simulated data.

We have seen in Figure 6.13, there is a brief period around 2.1 s when the predictions diverge from the data, but the driving currents eventually synchronize the two models. Table 6.4 shows the estimated and actual conductances in this experiment. For the 11 synapses actually present in the network, the estimated conductances are all nonzero and within 10–50% of the true values. Out of the 19 synapses not present in the network, 18 (94%) were estimated to have a maximal conductance less than 0.04 mS/cm^2. Most of the synapses erroneously identified to be present had small estimated conductances, with the exception of the inhibitory connection between Neurons 6 and 5. The activity of these neurons is negatively correlated, and thus, perhaps, it is unsurprising that a solution exists with a nonzero connection between them.

6.6 Comparison to GLM methods

Generalized linear models (GLMs) are a common tool for analyzing data that tends not to be Gaussian distributed. [Hardin and Hilbe, 2012] The most common version of this method for describing neural activity is the use of linear-nonlinear Poisson (LNP) cascade models [Gerwinn et. al. 2010, Stevenson et.al. 2008, Paninski, 2004] that treat neural activity as the sampling of an inhomogeneous Poisson process. Unlike the dynamical models described in this thesis, the LNP models generally do not contain any memory of prior states below the recorded event level - while spikes may alter the firing rates of other cells, the prior expected firing rate does not modify the future firing rate if it has not produced a spike)

$$\lambda(t) = \lim_{dt\to 0} \frac{1}{dt} P\big(N(t, t+dt) = 1\big) \tag{6.3}$$

where $P\big(N(t_1, t_2) = n\big)$ is the probability of n spikes between times t_1 and t_2. The number of spikes in a given interval is a Poisson process on the firing rate $\lambda(t)$ [Synder and Miller, 1991]:

$$P\big(N(t_0, t_1) = n\big) = \frac{1}{n!}\left[\int_{t_0}^{t_1} dt'\lambda(t')\right]^n \exp\left[-\int_{t_0}^{t_1} dt'\lambda(t')\right] \tag{6.4}$$

We are interested in finding the probability of some sequence of n spikes $\{t_j\}$. We know that for every interval $(t_{j-1}, t_j]$ there is going to be one spike. This leads to a product of smaller intervals:

$$P\big(\{t_j\}\big) = \lim_{dt\to 0} \prod_{j=1}^{n} P\big(N(t_{j-1} + dt, t_j) = 0\big) P\big(N(t_j, t_j + dt) = 1\big) \tag{6.5}$$

The product over the exponential component leads to an integral over the full range. The product over the first term works becomes a product of integrals over delta functions. Combining these leads to:

$$P\big(\{t_j\}\big) = \left[\prod_{j=1}^{n} \lambda(t_j)\right] \exp\left[-\int_{t_0}^{t_n} dt'\lambda(t')\right]. \tag{6.6}$$

For neural networks, the firing probability of any individual neuron will be some function of the external (to the organism) stimulus. Given some vector of

stimuli $\mathbf{s}(t)$ and the receptive field of the neuron $\mathbf{w}_s$, the activity of the neuron can be represented as:

$$\lambda(t) = f\big(\mathbf{s}(t)^T \mathbf{w}_s\big) \tag{6.7}$$

where f is some monotonically increasing positive function that grows at least linearly and at most exponentially [Paninski, 2004]. Note that $\mathbf{w}_s$ is not necessarily positive - a negative value indicates inhibition. This leads to a log likelihood for the spike times conditioned on the receptive field of:

$$\log p\Big(\{t_j\}|\mathbf{w}_s, \mathbf{s}(t)\Big) = \sum_{j=1}^{N} \left[\log \lambda(t_j)\right] - \int_0^T dt' \lambda(t') \tag{6.8}$$

$$= \sum_j \log\left[f(\mathbf{s}(t')^T \mathbf{w}_s)\right] - \int_0^T dt' f(\mathbf{s}(t')^T \mathbf{w}_s).$$

This formulation assumes that the firing rate is only a function of the stimulus to the system and is completely independent of both the current state of the neuron and the activity in the rest of the network. Obviously this is not true in a real network as the firing times will be correlated with the previous firing time of the current neuron due to the refractory period, bursting behavior, and other dynamical effects. The firing time will also be correlated (or anti correlated) with the firing times of other neurons due to the synaptic connections between them. This leads to a more complex activity vector:

$$\lambda(t) = f\Big(\mathbf{s}(t)^T \mathbf{w}_s + \Psi_h(t)^T \mathbf{w}_h\Big) \tag{6.9}$$

where $\Psi_h(t)$ is a time dependent linear filter that encodes the prior activity of the neurons in the network and $\mathbf{w}_h$ is analogous to the type and strength of the connections between them. This modifies the conditional probabilities:

$$\log p\Big(\{t_j^i\}|\mathbf{w}_s^i, \mathbf{w}_h^i\Big) = \sum_{i,j} \log \lambda^i\big(t_j^i\big) - \int_{t'=0}^T dt' \lambda^i(t') \tag{6.10}$$

$$= \sum_{j,i} \log f\Big(\mathbf{s}(t_j)^T \mathbf{w}_s^i + \Psi_h(t_j^i)^T \mathbf{w}_h^i\Big) - \int_{t'=0}^T dt' f\big(\mathbf{s}(t')\mathbf{w}_s^i + \Psi(t')^T \mathbf{w}_h^i\big)$$

where the superscript, 'i' is the index of the current neuron. Maximizing the first term is equivalent to maximizing the firing probability at the times of the actual spikes.

Minimizing the second term minimizes the chance of spike outside the measured times. For appropriate choices of f, maximizing this log likelihood allows for the estimation of receptive field and local connectivity parameters conditioned on a time series of spike times. [Paninski, 2004]

The dynamics of the neuron itself are encoded in the diagonal terms of the matrix Ψ. The choice of $\Psi(t)$ in LNP is essentially a linear combination of the local and current neuron activity. Just as we constructed stylized spikes to drive the neurons through action potentials and refractory periods to correspond to spike times, Ψ will contain a component that makes a spike extremely unlikely within 1-2 ms of the last spike, but more likely during the subsequent recovery. Thus in this region, Ψ will roughly track the state of the system in some similarly formulaic manner. The choice of stereotyped synaptic activity in response to measured somatic spikes creates a similar effect to the off diagonal terms in Ψ. Because of the post synaptic voltage dependence of the resulting current and potential filtering in the dendrites, the scalable activity pattern encoded in the off diagonal elements may need to be neuron state dependent - which may be better suited to the path integral method.

Physically the 'receptive field' of a neuron is a combination of the activity in projection neurons from prior subnetworks in the functional path and the connectivity of those projection neurons with neurons in the local network and is thus not directly analogous to the connectivity of the projection neurons. As of writing, we have not demonstrated the ability to determine the receptive field of neurons given some external stimulus simultaneous with unknown network architecture. In the limit where recurrent behavior is negligible we have demonstrated that we can estimate the physical currents acting on an identified cell through recordings of its voltage behavior (see Chap. 5).

At the beginning of this section, it was noted that the LNP methods are generally memoryless below the event level in the sense that the firing probabilities are only updated according to changes in stimuli - some of which are a function of prior spikes in neurons in the network. The methods used in this section for small networks of neurons could be argued to have the same issues. Given that measurements of spikes are required for the path integral method to create an estimate of a spike the path integral method does not appear to be an improvement in this regard. It is possible

that the stimulus dependent bifurcations in neuron behavior (and the subsequent effects on other neurons) cannot be estimated absent some measurements.

Acknowledgement

Figure 6.3 was reproduced by permission of authors from Meliza, C. D., M. Kostuk, H. Huang, A. Nogaret, D. Margoliash,and H. D. I. Abarbanel, "Estimating parameters and predicting membrane voltages with conductance-based neuron models" *Biol. Cybernetics*, **108** 495-516 (2014)

Chapter 7

Firing Rate Models

7.1 Lotka-Volterra

Spiking models - like the Hodgkin Huxley type models we have worked with until now, attempt to encode the full state of each neuron in attempt to accurately represent the underlying biology of the various voltage dependent ionic currents, changes in ionic concentrations, and other microscopic and fast properties. These models, while realistic, have the problem of complexity. The model describing the behavior of a real neuron from HVC has 12 dynamical variables and 69 parameters for a one compartment model of a single cell with no consideration of the various connections acting on the cell. Even discounting synapses, which scale in number roughly like the number of neurons squared, going up into any network of substantial size is going create computational stresses if we are limited to these precise models.

In order to reduce the models to a more simple form we must determine what features of the model are actually important to the function of the network - then create a reduced model that accurately encodes that information. In the case of some systems the exact temporal order of spikes provides little additional information over a time averaged firing rate of each neuron - or even group of neurons.

One such example system is the olfactory system in insects. Generally a set of odors will activate a set of sensory neurons - which in turn will stimulate a set of excitatory projection neurons and laterally inhibiting interneurons (mitral and granule cells). [Laurent, et al., 2001, Laurent, 2002] This lateral inhibition serves

to translate the noisy input into the sensory neurons into a spike sequence in the projection neurons that identifies the odor(s) present.

The insect olfactory system is of particular interest because the spike sequence identifying the odor is encoded in a stable temporal pattern. Rather than there being a 'mint' projection neuron or a 'strawberry' projection neuron, there appears to be a small number of projection neurons that exhibit a sequential series of activity switches that are unique to a particular odor. [Rabinovich et al, 2000]

This periodic switching is an indication of winnerless competition, essentially A inhibits B which inhibits C which inhibits A. provided A, B, and C are all stimulated to fire regularly this result means whenever A is active, B becomes less active, which in turn stops the inhibition on C, which activates and inhibits A, activating B. The temporal pattern of this cycling encodes the underlying sensory information using a only a few neurons in a way that is robust against noise. [Rabinovich et al, 2001]

Rather than use a spiking network, the firing rate of each projection neuron can be approximated [Fukai and Tanaka, 1997] by a Lotka-Volterra formulation [Lotka, 1910]:

$$\dot{a}_i(t) = a_i(t)\left[\sigma\big(\mathbf{a}(t), S_i^+(t)\big) - a_i(t) - \sum_{j\neq i}\rho_{ji}a_j(t) + S^-(t) + \eta^-(t)\right] \tag{7.1}$$

where $a_i(t)$ is the firing rate at time t of the i'th projection neuron.[1] σ is a smooth, sigmoidal function that qualitatively represents the non-linear threshold behavior of neurons in response to excitatory stimulus.

$$\sigma_i(S_i(t), \mathbf{a}(t)) = \tanh\left[5\left(S_i^+(t) - 0.4 + \eta^+(t) + g_E\sum_{j\neq i}a_j(t)\right)\right] \tag{7.2}$$

S(t) is the external stimulus to the system, broken up into an excitatory component in the sigmoidal function and inhibitory component outside. Because this stimulus is provided by the activity of the excitatory sensory neurons, this stimulus will generally be positive.

The nodes are coupled by way of a constant excitatory coupling to each other - ${g_E}^2$ in the sigmoidal function and through inhibitory connections defined according to

[1] because the 'neurons' in this model are assumed to project with both excitation and inhibition onto each other, each 'neuron' must contain mutually exclusive excitatory and inhibitory cells.

[2] or equivalently - the excitation is proportional to the average activity in the system

the matrix ρ_{ij}. Because the inhibition is actually from the activity of other neurons, this inhibition matrix should in principle be stimulus dependent. However, if we assume that the activity of these inhibitory cells are correlated with their respective projection neuron, we can drop this distinction. Thus the parameter values for the 'conductances' g_E and ρ_{ij} will not necessarily correspond exactly to the conductivity of the excitatory and inhibitory connections - but rather a combined effect of the excitatory and inhibitory components of the network.

To understand how this type of model can encode sensory information we must look to the structure of the attractors for different types of connections (ρ_{ij}) and stimuli. Regardless of stimuli, the Lotka Volterra model has fixed point at $\mathbf{a} = 0$. This fixed point is only stable as long as the sigmoidal function is negative; which is sensible as in the absence of stimuli can easily be translated into the absence of activity in the projection neurons. As the stimulus increases, there are additional fixed points at $a_i = \delta_{ij}\sigma_j$ corresponding to states without inhibition, the stability of which are dependent on the connectivity. Provided the inhibition is small enough, there will be an additional, stable fixed point when:

$$\left[\sigma\big(\mathbf{a}(t), S_i^+(t)\big) - a_i(t) - \sum_{j\neq i} \rho_{ji} a_j(t)\right] = 0 \tag{7.3}$$

Generally, this will only occur when all $\rho_{ij} < 1$. This case corresponds to a 'winner-shares' network where the output is simply a direct function of the input, independent of initial state. At $\rho_{ij} > 1$ Eq. 7.3 is only true when $a_i < 0$ for some or all i. Since the factor of a_i in front of this equation prevents a non-physical sign change in the activity, the dynamics never enters the basin of attraction of that fixed point. For this connection type, the fixed points at $a_i = \delta_{ij}\sigma_j$ become stable and the network is an example of 'winner takes all' as the first state to activate will suppress all other states. This leaves the cases where $\rho_{ij} \neq \rho_{ji}$ with one greater than 1 and the other less. In these cases there is the possibility of no stable fixed points and the dynamics will follow a complex limit cycle.

We turn to a three node case as a demonstration of this winnerless competition. Assume that the three nodes are identical ($\rho_{12} = \rho_{23} = \rho_{31} = \rho_a$) but non-symmetrical ($\rho_{12} \neq \rho_{21} = \rho_b$). For this exercise, we will set $\rho_a = 0.2$ and $\rho_b = 5.0$. As can be seen in Fig. 7.1, a series of constant stimuli results in heteroclinic switches between

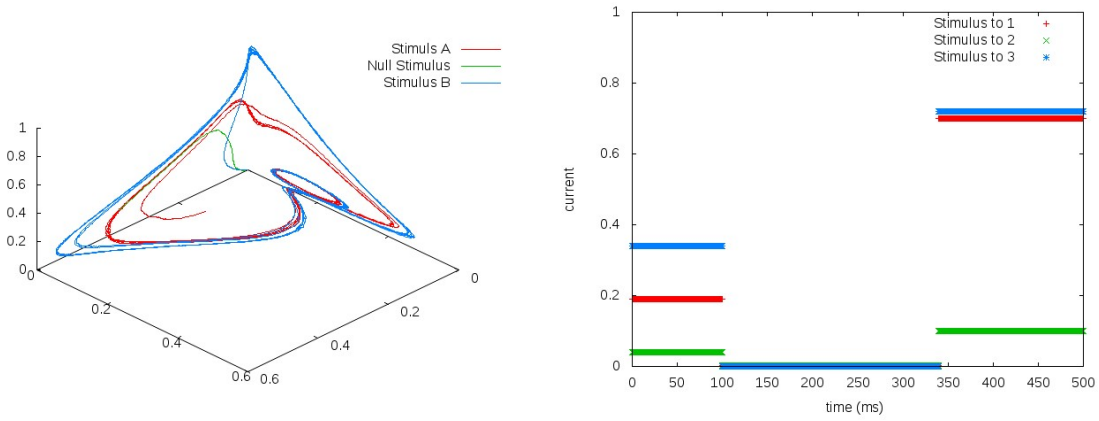

Figure 7.1: Phase plot of three state Lotka Volterra model with non-symmetric coupling (left) driven by a series of three constant stimuli (right). Differences in the shape of the resulting limit cycle (such as period doubling) allow for a large number of states to be conveyed in a temporal on a relatively small number of projection neurons. These patterns are robuts against noise as evidenced by the stability of these orbits despite an added 0.1 RMS Gaussian noise current.

fixed points. The order of the series of switches and the strength of the active state are robust against the noise 0.1 RMS Gaussian noise added to the system (20-30% of signal). The information about the stimulus is encoded in the strength and ordering of the switches between the different attractors. Changes in the stimulus, in addition to altering the strength of each switch will result in regions of period doubling due to orbits around the unstable fixed point at the winner share all fixed point. These stark changes in the structure of activity allow for a clear identification in a change in stimulus. [Rabinovich et al, 2001]

7.2 Twin Experiments on Lotka Volterra Networks

For our small spiking network in the previous chapter we demonstrated the ability to estimate the connectivity given full knowledge of the spiking times of every neuron and external stimulus to every neuron. This level of knowledge may not be realistic. While spike timing information allows for information about a larger fraction of the population, there is no guarantee that the source of a given spike can be attributed to a given neuron. Further, the stimulus to the network will often be some unknown and noisy translation to an external process. Given these limitations in

our measurements, can we extract model and stimulus information about a network of neurons in a similar fashion as we used to identify properties of an individual neuron.

When modeling individual neuron behavior we could treat the neurons in isolation from the network first then use the models developed to probe properties of the network. This worked in large part because in a single neuron in isolation we had both full control over the stimulus and an understanding of exactly how that stimulus appears in the model - as a current in a RC circuit. For the olfactory system the stimulus is presented to the measured projection neurons through excitatory signals from the sensory neurons. The activity of those sensory neurons are then a function of the their own internal dynamics and the local concentration of the chemicals the receptors on those sensory neurons are designed to detect. While that concentration can be controlled [Rabinovich et al, 2001], the stimulus that actually appears in the activity model of the projection neurons is an as yet unknown function of that concentration. The lack of knowledge about the actual stimulus along with a simultaneous lack of knowledge of the underlying connectivity of the network forces us to attempt to estimate the connectivity of the network simultaneously with the stimulus to the system.

To test if this is possible, we create a twin experiment of a three state Lotka Volterra model with winnerless competition. Activity data will be generated using a set of slowly varying stimuli with varying levels of added Guassian noise. The noise in the stimulus simulates the somewhat random timing of synaptic signals from the sensory neurons and random fluctuations in the concentration of odors at the receptors of the sensory neurons. Using measurements of each activity state, we use minAzero (See App. A) with Rm = 1 and Rf/dt^2 = 1e3 for each state. Similar to the stimulus estimates in neurons, we assume the stimulus acting on each projection neuron does not vary quickly in time. This is implemented by treating each stimulus as a state variable. By declaring that the time derivative of the stimulus is zero with a weak constraint ($Rf/dt^2 = 1$), we can allow the stimulus variable to vary freely but some regions are more free than others.

Fig. 7.2 shows the measured activity in each of the of the three nodes in the absence of noise (left) and the estimated currents for a variety of noise levels (right). Tab. 7.1 shows that estimates of the inhibition matrix ρ_{ij} and recurrant excitation

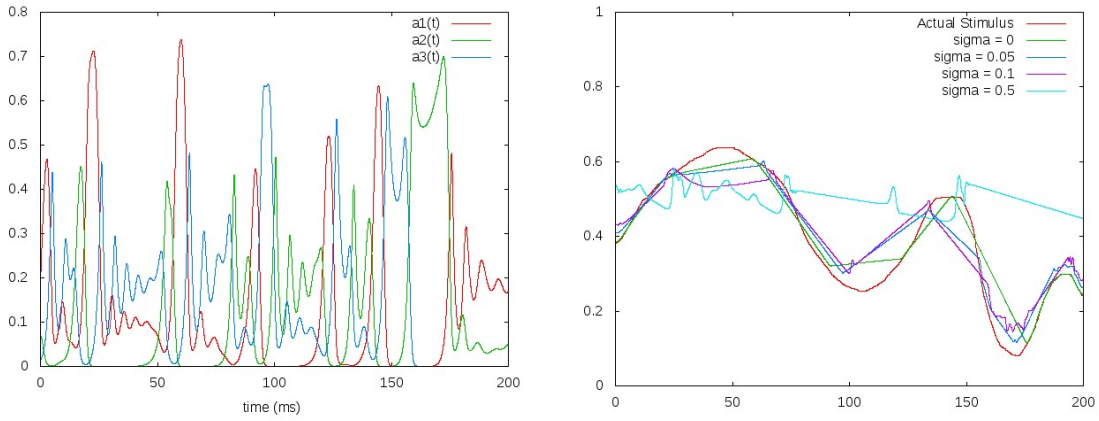

Figure 7.2: Plot of each of the three measured activity patterns (left) in the absence of noise in the stimulus. As the noise in the stimulus is increased, the quality of estimates of that stimulus decrease until it fails to track the general shape at a stimulus to noise ratio of about 1. The stepping behavior of the stimulus can be attributed in part to the periodic switching of states such that each current is only sampled during that state's turn in the sequence. Current estimations are made concurrently with estimates of the inhibitory connectivity matrix and mean excitatory conductance in Tab. 7.1

strength g_E. For little or no noise, we are able to *simultaneously* estimate the model parameters and stimulus driving the network. Of course little is somewhat relative - the noise levels at which this estimation gets acceptable results goes up to around $\sigma = 0.1$. When one notes that the maximum stimulus is around 0.7, the signal to noise ratios at which we can make good estimates are remarkable.

Table 7.1: Parameter estimates for the excitatory and inhibitory couplings in a three state Lotka Volterra with varying levels of Gaussian noise in the driving stimulus. Estimates are made simultaneously with estimates of the current based on measurements of the activity in each of the three nodes in Fig. 7.2.

σ	Act.	0	0.05	0.1	0.5
g_E	4.0	3.96	3.74	3.44	1.23
ρ_{21}	0.2	0.200	0.257	0.306	0.662
ρ_{31}	5.0	5.00	5.087	5.17	5.624
ρ_{32}	0.2	0.199	0.260	0.323	0.652
ρ_{12}	5.0	5.00	5.09	5.17	5.58
ρ_{13}	0.2	0.199	0.261	0.329	0.611
ρ_{23}	5.0	5.00	5.08	5.16	5.57

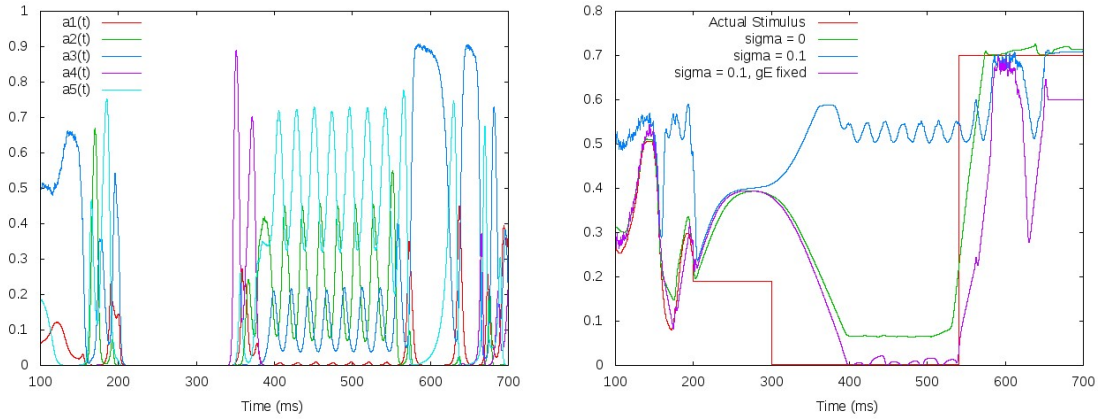

Figure 7.3: Plot of each of the measured activity patterns in a five state Lotka Volterra model (left) in the absence of noise in the stimulus. As the noise in the stimulus is increased, the quality of estimates of that stimulus (right - only stimulus to node 1 shown) decrease until it fails to track the general shape at a stimulus to noise ratio of about 1. The slow arc in the around the step occurs when $\mathbf{a}= 0$. The degeneracy at that fixed point prohibits us from making estimates of the stimulus in that region. Current estimations are made concurrently with estimates of the inhibitory connectivity matrix and mean excitatory conductance then again with the excitatory conductance fixed.

Table 7.2: Parameter estimates for five state Lotka Volterra network with unknown stimulus, inhibition (ρ_{ij}) and excitation (g_E). Parameter estimates are good until the noise in the current gets too large

σ	Act.	0	0.05	0.1		Act	0	0.05	0.1
g_E	1.5	1.37	1.117	0.017					
ρ_{21}	4.0	3.97	3.95	2.72	ρ_{43}	4.0	3.99	4.07	3.92
ρ_{31}	1.6	1.59	1.62	0.51	ρ_{53}	1.5	1.50	1.52	0.96
ρ_{41}	0.7	0.69	0.73	0	ρ_{14}	1.5	1.50	1.53	1.07
ρ_{51}	0.4	0.39	0.44	0	ρ_{24}	0.8	0.79	0.877	0
ρ_{12}	0.1	0.075	0.09	0	ρ_{34}	0.2	0.19	2.17	0.03
ρ_{32}	4.0	3.99	4.01	3.43	ρ_{54}	4.0	3.99	4.06	3.50
ρ_{42}	1.5	1.50	1.52	1	ρ_{15}	5.0	4.99	5.08	4.44
ρ_{52}	0.5	0.50	0.52	0	ρ_{25}	1.5	1.50	1.54	1.00
ρ_{13}	0.6	0.60	0.64	0	ρ_{35}	0.9	0.89	0.94	0.04
ρ_{23}	0.4	0.39	0.42	0	ρ_{45}	0.3	0.29	0.33	0

Table 7.3: In comparison with Tab. 7.2, fixing the excitation at the correct value provides sufficient additional information to estimate the inhibition parameters at larger levels of noise in the stimulus.

σ	Act.	0.1	0.2	0.5		Act	0.1	0.2	0.5
ρ_{21}	4.0	4.06	4.16	4.27	ρ_{43}	4.0	4.15	4.27	4.22
ρ_{31}	1.6	1.66	1.71	1.51	ρ_{53}	1.5	1.55	1.60	0
ρ_{41}	0.7	0.74	0.78	0.95	ρ_{14}	1.5	1.58	0.66	0
ρ_{51}	0.4	0.46	0.53	0.71	ρ_{24}	0.8	0.93	0.49	0
ρ_{12}	0.1	0.14	0.19	0.34	ρ_{34}	0.2	0.35	0	0
ρ_{32}	4.0	4.05	4.13	4.2	ρ_{54}	4.0	4.16	2.43	3.65
ρ_{42}	1.5	1.54	1.60	1.79	ρ_{15}	5.0	4.90	4.98	3.88
ρ_{52}	0.5	0.55	0.60	0.85	ρ_{25}	1.5	1.65	1.7	1.75
ρ_{13}	0.6	0.68	0.78	1.09	ρ_{35}	0.9	0	0.31	0
ρ_{23}	0.4	0.44	0.49	0.65	ρ_{45}	0.3	0.40	0.40	0.52

Appendix A

minAzero

minAzero is a python script used to write C++ code and compiler instructions using the IPOPT (Interior Point OPTimization) [Wächter & Biegler, 2006] libraries to estimate unmeasured states and parameters in dynamical systems with limited measurements. The scripts take a set of differential equations and state and parameter names provided by a text file "equations.txt" and returns a set of C++ files consisting of a set of constraints based on a discretized version of those differential equations. A second text file 'specs.txt' allows for changes in run specific quantities state and parameter bounds, as well as input files without the need to recompile.

These scripts and a few simple examples can be found at:

git@github.com:countfizix/minAzero

Installing Required Programs and Packages

Installing and running these codes currently requires a Linux distribution and sudo access. While it may be possible to install this setup on another Unix system (such as Mac) I have never attempted to do so.

Installing minAzero.py

Files:

- minAzero.py
  -Writes problem specific c++ file using strings from discAzero.py (xminAzero_nlp.cpp)
- discAzero.py
  -Discretizes equations and creates strings for Jacobian and Hessian Elements.
- makecppAzero.py
  -Writes master c++ file linking to IPOPT libraries (xminAzero_main.cpp)
- makehppAzero.py
  -Writes header file for x_nlp.cpp (xminAzero_nlp.hpp)
- makemakeAzero.py
  -Writes makefile for problem. Will need to be changed based on install location of IPOPT
- makeoptAzero.py
  -Writes settings file for IPOPT (x.opt)

where x is the problem name defined in equations.txt.

These python scripts require the sympy library. To install use sudo apt-get install sympy or download directly from sympy.org.

In order to link to your IPOPT libraries correctly, one line in makemakeAzero.py need to be modified. change line 59:

```
prefix = /home/mcserver/Desktop/Ipopt-3.11.7/build\n\
```

to the build directory for your IPOPT installation.

These files can be put in /usr/local/sbin for ease of use.

Installing IPOPT

Download

Get it here: https://projects.coin-or.org/Ipopt

- Download and unzip latest version of IPOPT
- As of right now this is 3.11.7 - Efficacy of installation instructions may degrade over time as packages are updated.
- Go into ThirdParty folder in the IPOPT directory then do the following commands.

```
$ cd Blas
$ ./get.Blas
$ cd ../Lapack
$ ./get.Lapack
$ cd ../ASL
$ ./get.ASL
$ cd ../Metis
$ ./get.Metis
```

- Get the HSL subroutines from http://hsl.rl.ac.uk/ipopt
- Note that there are two releases for HSL - you will want the more complete one that contains ma57, ma77, and ma97.
- While the freely available ma27 will work for many problems, the newer packages are faster, work on larger problems, and can use multi-core architecture.
- This will require filling out a form stating essentially that you are in academia and waiting a couple hours for a link to download.
- Unpack the resulting library into the ThirdParty folder such that the path is (IPOPT Path)/ThirdParty/HSL/coinhsl

Install

- Go to the IPOPT directory

```
$ mkdir build
$ cd build
$ ../configure
```

- Note that if you have lapack or blas installed previously you can use –with-lapack and –with-blas to link to those packages

- If something goes wrong refer here
  http://www.coin-or.org/Ipopt/documentation/node19.html#ExpertInstall

- Assuming everything worked:

```
$ make
$ make test
$ make install
```

Running the Code

minAzero uses two text documents (along with any needed data files) as input, equations.txt and specs.txt. Once these are filled

equations.txt contains information on the model and is used once for generating the needed cpp and hpp files for the run. The file should be written as described below in this order.

- The first line is the problem name, this name will be used to name the resulting executable.
- The second line tells minAzero how many dynamical variables, parameters, coupling terms, stimuli, functions, and measurements there are, in that order as a comma delimited list. It is essential that these numbers are accurate as minAzero uses this to know how many lines to read for each component of the code.
- A list of every differential equation.
- The measurement term of the cost function. A penalty term for coupling terms is suggested as any coupling to measurements is not present in physical systems.
- The names of all the variables. These must be the same as used in the differential equations and should be multiple letters/and or numbers such that variable name is contained in any other name or common function.
- The names of parameters, names of couplings, names of data, and names of stimuli, in that order. Again use fully unique names.
- Function names and number of arguments of that function separated by a comma. Use a function if there is some component of the dynamics with a removable singularity or other difficult numerical object that requires an alternative local definition.

- Functions will require an additional file 'myfunctions.cpp' containing the function definition along with its jacobian and hessian (an example of this is included)

specs.txt contains run specific information such as file names, variable bounds, and problem length. This file can be edited without recompiling the code.

- First line is the number of full steps the code will use. Because the code is compiled using a midpoint method, the actual problem length will double this plus one.
- Second line is the number of lines in each input file to skip. This allows for the code to start at any point in a long data set.
- Third line is double the time step of the data. Again since a midpoint method is used, the time step is for a whole step - which includes two points.
- If you wish to start at a non constant guess, you can put a 1 followed by a line with an initial condition file. This file should have one column for each state. If you do not want to include an initial condition file, use 0
- One line for each of the measured data file names. Each file should be a single column.
- One line for each of the stimulus data file names. Each file should be a single column.
- For each variable, the lower bound, upper bound, initial guess, and RF value separated by commas. The initial value is ignored if the initial conditions file is used, but a value must be included regardless.
- For each coupling term, a line with lower bound, upper bound, and initial value separated by commas followed by another line for lower bound, upper bound, and initial value for the derivative of the coupling term.
- For each parameter a lower bound, upper bound, and initial guess separated by commas

Once everything is filled out and all data files are present, you can run the python scripts:

```
$ minAzero.py
$ make
$ ./(problem_name)_cpp
```

If data files are missing or too short, the code will segfault. The outputs are **data.dat** containing all state variables at all times, and **param.dat** with parameter values.

Example equations.txt

```
# lines starting with # are ignored
# Problem Name
Colpitts
# nY,nP,nU,nI,nF,nM
3,3,1,0,0,1
# equations
yy+u00*(Data-xx)
-gam*(xx+zz)-qq*yy
eta*(yy+1-exp(-xx))
# measurement portion of Objective/Cost function
# the model portion is generated automatically
(Data-xx)*(Data-xx)+u00*u00
# variable names (nY)
xx
yy
zz
# parameter names (nP)
gam
qq
eta
# coupling names (nU)
# most minAzero formulated problems will not need a coupling term
# this term is essential for equality constrained dynamics
u00
# data names (nM)
Data
# stimuli names (nI)
# no stimuli in this problem
# functions (nF)
# list of functions and number of arguments - will require myfunctions.cpp
# with definitions of f(x) and its jacobian and hessian.
# alpha,4
```

Example specs.txt

```
# lines starting with # are ignored
# The problem length - actual length will be 4001 due to midpoint method
2000
# How much data to skip.
# In case you do not want to start at the beginning of the data file
10
# Time step - this is twice the time step of the data,
# since the data includes time and midpoints.
0.2
# Data File names - measurements
# each measurement needs its own file in its own row
testx.dat
#colscale1x.dat
# Data File names - stimuli
# each stimuli needs its own file in its own row
# No stimuli for this problem
# Boundary & initial conditions
# 0 for no initial data file, 1 for data file
# A data file must include values for all state variables at each time point.
0
# If above is 1, list name of data file next.  If 0, no entry needed.
#start.dat
# State Variable bounds:
# These are in the formats: lower bound, upper bound, initial guess, RF
# Boundary & initial conditions
# x
-100, 100, 20,1
# y
-100, 100, 10,1
# z
-100, 100, 1,1
# each coupling needs two entries, for u and du as there is no equation for u
# u00
# couplings are typically not needed in unconstrained problems
# but are essential in equality constraint problems
0,1,0
-10,10,0
# each parameter needs upper bound, lower bound, initial guess
# p1
0, 100, 0.016
# p2
0, 100, 0.14
```

```
# p3
0, 100, 1.26
```

Appendix B

Spike Timing

In describing the form of the action potential we only considered the role the extracellular medium plays in determining the reversal potential of the various ionic species. For point models this is appropriate provided the flux of ions is small compared to the size of the reservoir and thus a constant reversal potential. For real neurons with physical extent the behavior of the neuron leads to effects in the extracellular medium that can be measured.

The propagation of action potential acts as a front of depolarizing current followed by front of repolarizing current. The propagation of this wave creates a series of internal, effectively static[1] electric fields in the cell as shown in Fig B.1. Because charge is locally conserved, this internal field must correspond to an equal and opposite external field.

To calculate the magnitude of this effect we first assume that the neuron is an infinitely long cylinder of radius a and that the voltage satisfies Laplace's equation. The resulting symmetry gives a differential voltage per unit length of:

$$dV_{ext} = \frac{i_m(x)dx}{4\pi\sigma_{ext}r} \tag{B.1}$$

where $i_m(x)$ is the membrane current per unit length as a function of x and σ_{ext} is the conductance of the extracellular medium. Recalling from the discussion on the cable equation that because of current conservation the membrane current is proportional to the internal divergence of the voltage (Eq. 4.2).

[1] no magnetism

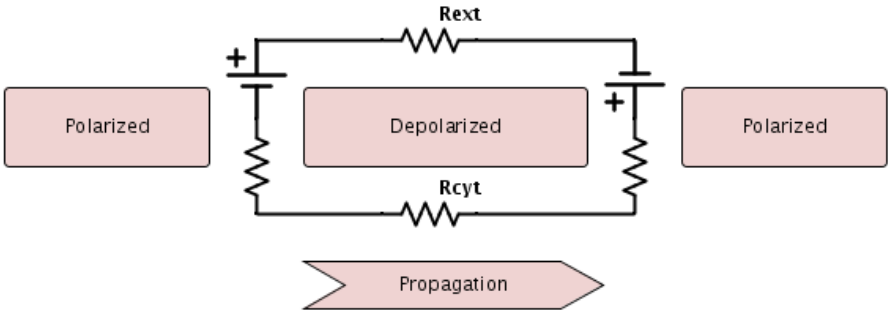

Figure B.1: Rough circuit diagram for the instantaneous currents during the propagation of an action potential. A front of depolarizing current followed by a repolarizing front creates an internal potential in the cytoplasm. Conservation of charge requires that the sum of all currents in the loop be zero - indicating the presence of an external current opposite the propagation of the spike. This circuit creates a time varying dipole field in the extracellular medium that can be detected with an extracellular probe.

$$i_m = \pi a^2 \sigma_{cyt} \frac{\partial^2 V_m}{\partial x^2} \tag{B.2}$$

giving

$$V_{ext}(\mathbf{r}, t) = \frac{a^2 \sigma_{cyt}}{4\sigma_{ext}} \int \frac{1}{r} \frac{\partial^2 V_m}{\partial x^2} dx \tag{B.3}$$

where σ_{cyt} conductance of the cytoplasm.[2] Integrating by parts x gives:

$$V_{ext}(\mathbf{r}, t) = \frac{a^2 \sigma_{cyt}}{4\sigma_{ext}} \int \frac{\partial V_m(x,t)}{\partial x} \nabla \frac{1}{r} \cdot \hat{\mathbf{x}} dx \tag{B.4}$$

Without knowing $V_m(x,t)$ precisely we note that this integral is largest when the gradient of the voltage is large. Since the spatial gradients come from the propagation of a fixed waveform down the cell, these spatial gradients are equivalent to temporal ones at a fixed point as in our point model. Because of their shape, action potentials well present a large extracellular signal due to the closely spaced regions of rapid voltage changes - orders of magnitude greater than the subthreshold behavior. Note that the dot product of the gradient with $\hat{\mathbf{x}}$ causes a rise in voltage followed by a fall in voltage to be additive, the means the peak value for the voltage signal in the extracellular medium corresponds directly to the peak of the action potential.

[2]Note that in the previous discussion on the cable equation we assumed the extracellular potential was constant. Replacing V_m with $V_m + V_{ext}(r)$ in equation B.2 allows for the potential to be solved iteratively should additional precision be needed.

Bibliography

[Abarbanel, 1996] H. D. I Abarbanel, *Analysis of Observed Chaotic Data*, Springer-Verlag, 1996

[Abarbanel, 2013] H. D. I Abarbanel, *Predicting the Future: Completing Models of Observed Complex Systems*, Springer-Verlag, Spring, 2013.

[Abarbanel, et al., 2009] H. D. I. Abarbanel, D. R. Creveling, R. Farsian, and M. Kostuk, "Dynamical State and Parameter Estimation," *SIAM J. Appl. Dyn. Syst.* **8**, 1341-1381 (2009) .

[Anderson et al, 2014] D. Anderson, B. Ermentrout, and P. Thomas, "Stochastic Representations of Ion Channel Kinetics and Exact Stochastic Simulation of Neuronal Dynamics," arXiv:1402.2584 (2014)

[Buzsaki, 2004] G. Buzsáki, " Large-scale recording of neuronal ensembles," *Nat Neurosci.* **7**, 446-51, (2004).

[Carrassi and Vannitsem, 2011] A. Carrassi and S. Vannitsem, "State and parameter estimation with the extended Kalman filter: an alternative formulation of the model error dynamics," *Q. J. R. Meteorol. Soc.* **137**, 435-451, (2011)

[Certik, 2006] O., Certik, "SymPy Library for Symbolic Mathematics," *Technical report, http://code.google.com/p/SymPy/*, since 2006.

[Chapin, 2004] J. Chapin, "Using multi-neuron population recordings for neural prosthetics," *Nature Neuroscience*, **7**, 452-455, (2004)

[Chou, 1996] M. Chou and Y. Lin, "Exotic Dynamic Behavior of the Forced FitzHugh-Nagumo Equations," *Computers Math. Applic.* **32**:10 109-124 (1996)

[Cowan, Südhof, & Stevens, 2001] W. Maxwell Cowan, Thomas C. Südhof, and Charles F. Stevens, *Synapses*, Johns Hopkins University Press, Baltimore (2001)

[Creveling, et al., 2008] Creveling, D. R., P. E. Gill and H. D. I. Abarbanel, "State and parameter estimation in nonlinear systems as an optimal tracking problem," *Physics Letters A* **372**, 2640-2644, (2008).

[Daou, et al, 2013] A. Daou, M. T. Ross, F. Johnson, R. L. Hyson, and R. Bertram, "Electrophysiological characterization and computational models of HVC neurons in the zebra finch" *Journal of Neurophysiology*, **10** 1227-45 (2013)

[Davie, et. al. 2001] J. Davie, M. Kole, J. Letzkus, E. Rancz, N. Spruston, G. Stuart, and M. Häusser "Dendritic patch clamp recording", *Nature Protocols* **1** 1235-1247 (2006)

[Denk et al., 1990] Denk, Winfried, James H. Strickler, and Watt W. Webb. "Two-photon laser scanning fluorescence microscopy." *Science* **248** 73-76 (1990): 73-76.

[Destexhe, Mainen, & Sejnowski, 1994] A. Destexhe, Mainen Z. and Sejnowski T. "Synthesis of Models for Excitable Membranes, Synaptic Transmission and Neuromodulation Using a Common Kinetic Formalism." *Journal of Computational Neuroscience***1**, 195-230 (1994).

[Destexhe & Sejnowski, 2001] Alain Destexhe and Terry Sejnowski, *Thalamocortical Assemblies: How Ion Channels, Single Neurons and Large-Scale Networks Organize Sleep Oscillations*, Oxford University Press, USA 2001 ISBN: 0198524250 / 0-19-852425-02001.

[Fano, 1961] Fano, R. M., *Transmission of Information: A Statistical Theory of Communication*, Wiley, New York, (1961).

[FitzHugh, 1955] R. FitzHugh, "Mathematical models of threshold phenomena in the nerve membrane," *Bull. Math. Biophysics*, **17**:4 257-278 (1955)

[Fokker, 1914] A. D. Fokker, Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld, *Ann. Phys.* 348 **43(4)**, 810820 (1914).

[Fukai and Tanaka, 1997] T. Fukai and S. Tanaka, "A Simple Neural Network Exhibiting Selective Activation of Neuronal Ensembles: From Winner-Take-All to Winner-Share-All," *Neural Computation*, **9** 77-97 (1997)

[Gerwinn et. al. 2010] S. Gerwinn, J. H. Macke, and M. Bethge "Bayesian inference for generalized linear models for spiking neurons", *Front. Comput. Neurosci.*, **4**:12. doi: 10.3389/fncom.2010.00012

[Gill, et al., 2005] Gill, P.E., W. Murray, and M.A. Saunders, "SNOPT: An SQP Algorithm for Large-scale Constrained Optimization," *SIAM Review* **47(1)**, 99-131 (2005).

[Gold, et al., 2007] Gold, C, Henze, DA, Koch, C. "Using extracellular action potential recordings to constrain compartmental models," J Comput Neurosci 23(1): 39–58. doi:10.1007/s10827-006-0018-2.

[Golding et al. 2002] N. L. Golding, N. P. Staff, and N. Spruston, "Dendritic spikes as a mechanism for cooperative long-term potentiation," *Nature*,**418**, 326-331 (2002)

[Gray, 1959] Gray, E. G. "Axo-Somatic and Axo-dendritic synapses of the cerebral cortex: An electron microscope study," *J Anat*, **93(4)**, 420-433, (1959)

[Hamill, et al., 1981] Hamill, O. P., Marty, A., Neher, E., Sakmann, B., Sigworth, F. J., "Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches", *Pflugers Archiv: European Journal of Physiology* **391**, 85-100 (1981) doi:10.1007/BF00656997.

[Hardin and Hilbe, 2012] Hardin, James W. and Hilbe, Joseph M. "Generalized linear models and extensions" Stata Press, 2012 ISBN 9781597181051

[Hastings, 1970] Hastings W. K., "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika* **57**, 97-109 (1970)

[Häusser et al, 2000] M. H ausser, N. Spruston, and G. Stuart, "Diversity and Dynamics of Dendritic Signaling,"*Science*, **290**(5492), 739-744 (2000)

[Hille, 2001] Hille, Bertil, "Ion channels of excitable membranes, 3rd Ed" Sinauer Associates, 2001 ISBN 0-87893-321-2

[Hodgkin and Huxley, 1952] Huxley AF Hodgkin AL. "A quantitative description of membrane current and its application to conduction and excitation in nerve." *J Physiol*, **117**, 1952.

[Johnston and Wu, 1995] Johnston, Daniel and Samuel M. S. Wu, *Foundations of Cellular Neurophysiology,* The MIT Press 1995 ISBN: 0 262 10053 3

[Julier and Uhmann, 1997] S. Julier and J. K. Uhlmann. "A new extension of the Kalman filter to nonlinear systems." *Int. symp. aerospace/defense sensing, simul. and controls.* **3(26)** (1997).

[Knowlton et. al. 2014] C. Knowlton, C.D. Meliza, D. Margoliash, and H.D.I. Abarbanel, "Dynamical estimation of neuron and network properties III: network analysis using spike times" *Biological Cybernetics* **108**:3 261-273 (2014)

[Kostuk, et al., 2012] Kostuk, M., B. Toth, C. D. Meliza, H. D. I. Abarbanel, and D. Margoliash, "Dynamical Estimation of Neuron and Network Properties II: Monte Carlo Methods," *Biol. Cybernetics*, **106**, 155-167 (2012).

[Kolmogorov A., 1931] Kolmogorov A., "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung," *Mathematische Annalen* **104**, 415-458 (1931)

[Langevin, 1908] Langevin, P. (1908). "Sur la thorie du mouvement brownien [On the Theory of Brownian Motion]". C. R. Acad. Sci. (Paris) 146: 530533. ; reviewed by D. S. Lemons & A. Gythiel: Paul Langevins 1908 paper "On the Theory of Brownian Motion" [...], Am. J. Phys. 65, 1079 (1997)

[Laplace, 1774] P. S. Laplace, "Memoir on the probability of causes of events," *Mémoires de Mathmatique et de Physique*, **16** (1774) (English translation by S. M. Stigler *Statist. Sci.* **1** 364378 (1986)).

[Laurent, et al., 2001] G. Laurent, M. Stopfer, R.W. Friedrich, M.I. Rabinovich, A. Volkovskii, and H.D.I. Abarbanel, "Odor Encoding as an Active Dynamical Process: Experiments, Computation, and Theory," *Annual Reviews of Neuroscience* **24**, 263-297 (2001).

[Laurent, 2002] G. Laurent, "Olfactory network dynamics and the coding of multidimensional signals," *Nature Reviews Neuroscience*, **3**, 884-895 (2002)

[Lorenz, 1963] Lorenz, Edward N., 1963: Deterministic Nonperiodic Flow. *J. Atmos. Sci.*, **20**, 130141

[Lotka, 1910] Lotka, A.J., "Contribution to the Theory of Periodic Reaction", *J. Phys. Chem.*, **14** (3), pp 271274 (1910)

[Meliza, et al., 2014] Meliza, C. D., M. Kostuk, H. Huang, A. Nogaret, D. Margoliash, and H. D. I. Abarbanel, "Estimating parameters and predicting membrane voltages with conductance-based neuron models" *Biol. Cybernetics*, **108** 495-516 (2014)

[Metropolis, et al., 1953] Metropolis, Nicholas and Rosenbluth, Arianna W. and Rosenbluth, Marshall N. and Teller, Augusta H. and Teller, Edward, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.* **21**, 1087-1092 (1953).

[Neher, et al., 1981] Nehar E Sakmann B Sigworth FJ Hamill OP, Marty A. Improved patch clamp techniques for high-resolution current recording from cells and cellfree membrane patches. Pugers Arch, 391(2):85-100, (1981)

[Nottenbohm, 2005] F. Nottenbohm "The Neural Basis of Birdsong" *PLoS Biol.*, **3**(5):e164 (2005)

[Oseledets, 1968] V.I. Oseledets, A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, Trans. Moscow Math. Soc. 19 (1968), 197-231. Moscov.Mat.Obsch.19, 179-210. (1968)

[Paninski, 2004] L. Paniniski, "Maximum likelihood estimation of cascade point-process neural encoding models" *Network: Comput. Neural Syst.*, **15**, 243-262 (2004)

[Poincaré, 1892] Poincar, H. (1892), "Sur les courbes dfinies par une quation diffrentielle", *Oeuvres* **1**, Paris

[Quinn & Abarbanel, 2010] J. C. Quinn and H. D. I. Abarbanel, "State and parameter estimation using Monte Carlo evaluation of path integrals," *Quarterly Journal of the Royal Meteorological Society* **136**, 1855-1867 DOI:10.1002/qj.690 (2010).

[Rabinovich et al, 2000] M. Rabinovich, R. Huerta, A. Volkovskii, H.D.I. Abarbanel, M. Stopfer, and G. Laurent, "Dynamical coding of sensory information with competitive networks," *J. Physiology-Paris*, **94** 465-471 (2000)

[Rabinovich et al, 2001] M. Rabinovich, A. Volkovskii, P. Lecanda, R. Huerta, H.D.I. Abarbanel, and G. Laurent, "Dynamical Encoding by Networks of Competing Neuron Groups: Winnerless Competition," *Phys Rev Letters*, **87** (2001)

[Rabinovich, 2007] T. Nowotny and M. Rabinovich. "Dynamical Origin of Independent Spiking and Bursting Activity in Neural Microcircuits" *Phys Rev Letters*, **98** (2007)

[Ramirez, 2004] JM Ramirez, AK Tryba, and F. Penã "Pacemaker neurons and neuronal networks: an integrative view" *Curr. Opin. Neurobiol.*,**14**(6), 665-74, (2004)

[Roth, 2001] A. Roth and M. Häusser "Compartmental models of rat cerebellar Purkinje cells based on simultaneous somatic and dendritic patch-clamp recordings," *Journal of Physiology*,**535**, 445-472 (2001)

[Smetters, Majewska, and Yuste, 1999] Diana Smetters, Ania Majewska, and Rafael Yuste. Detecting action potentials in neuronal populations with calcium imaging. Methods, 18(2):215- 221, 1999.

[Synder and Miller, 1991] Donald L. Snyder and Michael I. Miller, "Random Point Processes in Time and Space", Springer-Verlag 1991 ISBN 0-387-97577-2

[Stevenson et.al. 2008] I. H. Stevenson, J. M. Rebesco, L. E. Miller, and K. P. Körding "Inferring functional connections between neurons", *Curr. Op. Neurobio.* **18**, 582-588 (2008)

[Toth, 2010] B. A. Toth, "Python Scripting for Dynamical Parameter Estimation in IPOPT," *SIAG/OPT Views-and-News***21:1**, 1-8 (2010).

[Toth, et al., 2011] B. A. Toth, Kostuk, M., C. D. Meliza, D. Margoliash, and H. D. I. Abarbanel,"Dynamical Estimation of Neuron and Network Properties I: Variational Methods," *Biological Cybernetics*, Biol Cybern DOI 10.1007/s00422-011-0459-1, 1 - 21, (2011).

[Wächter & Biegler, 2006] A. *Wä*chter and L. T. Biegler, "On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming," *Mathematical Programming* **106**, 25-57 (2006).

[Zhang & Trussell, 1994] Zhang, Su and Trussell, Laurence, "Voltage clamp analysis of excitatory synaptic transmission in the avian nucleus magnocellularis" *Journal of Physiology*, **480,1** 123-136 (1994)