Channel Noise And Firing Irregularity In Hybrid Markov Models Of The Morris-Lecar Neuron

by

Casey Bennett

Submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mathematics, Applied Mathematics, and Statistics

Case Western Reserve University

Janurary, 2016

Case Western Reserve University
School Of Graduate Studies

We hereby approve the thesis of

Casey Bennett

candidate for the degree of Master Of Science

Signed: _____

Dr. Peter J. Thomas, Committee Chair, Advisor

_____

Dr. David Friel, External Faculty, Advisor

_____

Dr. Wojbor Woyczynski, Professor of Mathematics

_____

Dr. Mark Meckes, Professor of Mathematics

Date: August 3, 2015

We also certify that written approval has been obtained for any proprietary material contained therein

To My Father

Whose enthusiastic and intuitive grasp of science shaped my mind at young age.

An excerpt from *Brewer's Dictionary Of Phrase and Fable* (1898):

Hair of the Dog that Bit You (A). *Similia similibus curantur.* In Scotland it is a popular belief that a few hairs of the dog that bit you applied to the wound will prevent evil consequences. Applied to *drinks*, it means, if overnight you have indulged too freely, take a glass of the same wine next morning to soothe the nerves. "If this dog do you bite, soon as out of your bed, take a hair of the tail in the morning."

> Take the hair, it's well written,
> Of the dog by which you're bitten;
> Work off one wine by his brother,
> And one labour with another ...
> Cook with cook, and strife with strife:
> Business with business, wife with wife.
> *Athenus (ascribed to Aristophanes).*

> There was a man, and he was wise,
> Who fell into a bramble-bush
> And scratched out both his eyes;
> And when his eyes were out, he then
> Jumped into the bramble-bush
> And scratched them in again.

Contents

Nomenclature vii

List of Figures viii

Abstract ix

1 Channel Noise and Spiking Variability 1
- 1.1 Functional Role of Neural Variabillity . . . 1
- 1.2 Foundational Results of Computational Neuroscience . . . 4
- 1.3 Experimental and Computational Studies Which Illuminate the Role of Channel Noise in Neural Dynamics . . . 6
  - 1.3.1 Experimental . . . 7
  - 1.3.2 Computational . . . 8
  - 1.3.3 Beyond Stochastic Hodgkin-Huxley Models . . . 9
- 1.4 Analytical Studies of Neural Models . . . 10
- 1.5 Stochastic Ion Channel Kinetics Schemes: Connections and Their Consequences . . . 11
  - 1.5.1 Simulating Discrete Markov Models of the Hodgkin-Huxley Neuron . . . 12
  - 1.5.2 The Emergence of a Langevin Approximation . . . 14
  - 1.5.3 Conflict Over Fox and Lu's Langevin Approximations . . . 16
  - 1.5.4 Resolution, Effective Langevin Approximations, and Reflections on Channel Noise Models . . . 17
- 1.6 Contributions of This Project To Understanding Channel Noise and Spiking Variability . . . 20

2 Overview of Central Model and Motivations 24
- 2.1 Definition and Dynamics of the Exact Hybrid Markov Model of the Morris-Lecar Neuron . . . 25
  - 2.1.1 A Simple Model for Stochastic Currents . . . 25
  - 2.1.2 The Exact Hybrid Markov Model of the Morris-Lecar Neuron 26
  - 2.1.3 Dynamics, Spiking, and the Interspike Interval Distribution . . 30
- 2.2 Motivating Questions . . . 33
- 2.3 Methods . . . 34

3 Investigating The Channel Number ISI Surface **37**
3.1 Channel Configuration Space and the ISI Surface 37
3.1.1 Overview of Stochastic Approximation Schemes Across Channel Configuration Space 39
3.2 Scaling Across the ISI Surface 40
3.2.1 Intensive vs Extensive Expressions of the Current Equation . . 40
3.2.2 Stochastic Ion Channel Kinetics as Chemical Reaction Networks: Convergence of Markov Chain Models to Deterministic Rate Equations 43
3.2.3 Scaling Framework 46
3.2.4 Stochastic-Deterministic Limits 48
3.2.5 Scaling in the Literature 49

4 The Exact Poisson Region **50**
4.1 Exact Simulation Algorithm 51
4.1.1 Pathwise Representation of Unit-Rate, Channel-State Processes 52
4.1.2 Pathwise Representation of Channel-State Processes 54
4.1.3 Simulating Pathwise Representations of Channel-State Processes 55
4.2 Piecewise Constant Propensity Approximation (PCPA) 61
4.3 Numerical Convergence Between Exact and PCPA Algorithms 66

5 The Gaussian Region **71**
5.1 Gaussian Approximation of PCPA Algorithm 73
5.2 SDE Simulation Algorithm 77
5.3 Numerical Convergence of PCPA and Stochastic Differential Equation Model 80

6 The Weak Noise Region **87**
6.0.1 Introduction to Neural Oscillators and Phase 88
6.1 The Itō Approach to Weak Stochastic Perturbations 89
6.2 Weak Deterministic Perturbations of Neural Oscillators 92
6.2.1 Weak Individual Perturbations 93
6.2.2 Weak Finite Perturbations 96
6.2.3 Weak Continuous Perturbations 97
6.3 Weak Stochastic Perturbations of Neural Oscillators 100
6.3.1 Weak Stochastic Individual Perturbations 100
6.3.2 Weak Stochastic Finite Perturbations 102
6.3.3 Weak Stochastic Continuous Perturbations 104
6.4 Weak Stochastic Morris-Lecar Perturbations of Neural Oscillators . . 109
6.4.1 Interspike Interval Distribution of Neural Oscillator Subject to Stochastic Morris-Lecar Perturbations 110
6.4.2 Numerical Convergence of SDE Model to Weak Perturbation Plane 112

7 Statistical Convergence Between Models **114**
7.1 Estimating Error of Sample Mean, Variance, and CV 115
7.1.1 Notation and the Central Limit Theorem 115
7.1.2 Error of Sample Mean in Terms of Variance 116
7.1.3 Error of Sample Variance in Terms of Fourth Moment and Variance 117
7.1.4 Error of Sample CV in Terms of Kurtosis and CV 120
7.2 Estimating Error of Sample Mean, Variance, and CV in Terms of Measurable Quantities 122
7.2.1 Error of Sample Mean In Terms of Sample Variance 122
7.2.2 Error of Sample Variance In Terms of Fourth Sample Moment and Sample Variance 123
7.2.3 Error of Sample CV in Terms of Sample Kurtosis and CV . . 125
7.3 Numerical Validation of Asymptotic Approximations 126
7.3.1 Method of Validation 126
7.3.2 Verification and Scope of Standard Error Approximation . . . 128
7.4 Convergence of ISI Statistics 131

Nomenclature

ρ_{M}	Density of M Channels $\left(\text{Parameter } \frac{N}{\mu m^2}\right)$
ρ_{N}	Density of N Channels $\left(\text{Parameter } \frac{N}{\mu m^2}\right)$
$\rho_{\mathrm{M_F}}$	Fixed Density of M Channels $\left(60\frac{N}{\mu m^2}\right)$
$\rho_{\mathrm{N_F}}$	Fixed Density of N Channels $\left(18\frac{N}{\mu m^2}\right)$
A	Cell (or Membrane Patch) Area (Parameter μm^2)
A_{F}	Fixed Cell (or Membrane Patch) Area ($1\mu m^2$)
C	Specific Capacitance of Membrane $\left(20\frac{\mu F}{\mu m^2}\right)$
G_{L}	Maximal Specific Conductance to Leak Currents $\left(2\frac{\mu S}{\mu m^2}\right)$
G_{M}	Maximal Specific Conductance to M Currents $\left(4.4\frac{\mu S}{\mu m^2}\right)$
G_{N}	Maximal Specific Conductance to N Currents $\left(8\frac{\mu S}{\mu m^2}\right)$
I_{app}	Specific Applied Current $\left(100\frac{nA}{\mu m^2}\right)$
N_{M}	Total Number of M-Type Channels (Analogous to Calcium) (Parameter N)
$n_{\mathrm{M}}(t)$	Number of open M Channels (Variable N)
N_{N}	Total Number of N-Type Channels (Analogous to Potassium) (Parameter N)
$n_{\mathrm{N}}(t)$	Number of open N Channels (Variable N)
V	Membrane Voltage (Variable mV)
V_{L}	Leak Reversal Potential (-60 mV)
V_{M}	M-Type Reversal Potential (120 mV)
V_{N}	N-Type Reversal Potential (-84 mV)

List of Figures

1.1 Comparison of Model Runtimes And Validity As A Function of Channel Number 21

2.1 Rate Functions of the Morris-Lecar Neuron 28
2.2 Causal Structure of Neural Dynamics 29
2.3 Measured Spiking Activity of Purkinje Cell 30
2.4 State Space Dynamics of Hybrid Markov Model 31
2.5 Numerical Solutions of Central Model Which Establish Need For A Dual Threshold Spike Finder 32
2.6 Effect of Leaner Mutation on ISI Variability 33

4.1 Traces of Exact Algorithm 60
4.2 Runtime of Exact Channel Number Tracking Algorithm 61
4.3 Traces of PCPA Algorithm 63
4.4 Convergence of Exact Algorithm to PCPA 69
4.5 ISI Statistics From Exact and PCPA Models Do Not Converge 70

5.1 Runtime of Piecewise Constant Propensity Approximation and Stochastic Differential Equation Algorithms 72
5.2 Traces of SDE Algorithm (100 Channels) 78
5.3 Traces of SDE Algorithm (1000 Channels) 79
5.4 Convergence of PCPA Model to SDE Model 83
5.5 ISI Statistics From PCPA and SDE Models Converge 84
5.6 Survey of SDE Surfaces (Up to 10000 Channels) 85
5.7 Survey of SDE Surfaces (Up to 3000 Channels) 86

6.1 SDE Simulations Compared to Weak Noise Plane 113

7.1 Numerical Validation of Asymptotic Approximation of Standard Error 130

Channel Noise And Firing Irregularity In Hybrid Markov Models Of The Morris-Lecar Neuron

Abstract

by

Casey Bennett

Using a stochastic version of the Morris-Lecar model neuron, a scaling method is introduced in which the ODE that propagates voltage is invariant, but the underlying Markov chain which controls the discrete channel states converges to progressively simpler stochastic descriptions. The relationship between the underlying stochastic description of channel states can then be directly related to the induced interspike interval (ISI) variability as the system is scaled. Specifically, an exact Markov chain model, a piecewise constant propensity approximation to the exact model, and a Langevin approximation are derived. For large systems the ISI variance is found in terms of channel specific phase response curves. Through the development of powerful numerical implementations on a high performance computing cluster and efficient error estimation techniques, convergence of each model to the others is systematically evaluated, so that given any channel configuration and error tolerance, the fastest method can be easily identified.

Chapter 1

Channel Noise and Spiking Variability

1.1 Functional Role of Neural Variabillity

The contrast between the precision of the brain and its fundamental processing unit, the neuron, is one of the foremost mysteries of cognition. How can a skilled archer or grandmaster chess player quickly and consistently perform precise calculations with neurons that often produce radically different outputs given the same input? Von Neumann, in 1956, was the first to confront the problem with his groundbreaking paper, *Probabilistic logics and the synthesis of reliable organisms from unreliable components* [61]. While Von Neumann's work outlined potential methods of redundancy which could be used in the brain to overcome its unreliable constituents, he described the approach to be largely insufficient. Though the mystery persisted, shortly before his death, Von Neumann ushered in a new perspective on cognition through an unfinished monograph, *The Computer and the Brain* [62]: the fundamental principles which underlie cognition are not deterministic or exact, but random and statistical. Six years after his death this message was realized in a pioneering paper

Reliable Computation in the Presence of Noise by Cowan and Winograd [68], which advanced the understanding of neuronal computation by applying information theory to a re-envisioned nervous system composed of electrical components which generate and respond to noisy signals. The crucial insight established by Von Neumann and realized by Cowan and Winograd is that the function of the brain must be understood not just through the logical structure of its components, but also through the stochastic nature of its signals. This project focuses the insight of Cowan, Winograd, and Von Neumann into an analysis of signals which result from stochastic neural models to understand how certain mutations alter the dynamics, and critically disrupt the function of the nervous system. To contextualize future work, this effort will begin with a review of modern research on the functional role of neural variability in cognition.

Drawing intuition from digital computers, which perform skilled tasks using (approximately) deterministic subcomponents, it is natural to think of the stochastic nature of neurons as a side effect of biological machinery which must be overcome. However, a growing body of evidence suggests that neural variability may in fact be a prerequisite feature of the nervous system which enables computation. It appears, for instance, that nature has structured the nervous system to utilize random electrical fluctuations as a tool to amplify weak periodic signals [41]. Experiments on crayfish mechanoreceptors [14], cricket cercal sensory systems [39], and human tactile sensation [11] have all shown "stochastic resonance", or increased signal detection in the presence of noise. Similarly, computational models of the Hodgkin-Huxley neuron showed that periodic signals hidden in noise are amplified or reduced based on cell area [41, 69, 38, 29, 53]. One could interpret this as a feature of the nervous system which enables nature to "choose" a cell size which permits or denies the passage of certain frequencies or amplitudes [53].

Other useful consequences of neural variability have also been suggested. The com-

putation of uncertain information may be performed in the brain by using Bayesian inference to set up an appropriate probability distribution, and then the stochastic properties of neurons to perform a Markov Chain Monte Carlo type sampling method [5]. Another more general interpretation of the functional role of neural variability was offered by [54]: Information is transmitted through neurons in the form of spike time and spike rate. Neural noise does not have a significant effect on spike times originating from strongly correlated, instantaneous inputs. However, weak, uncorrelated inputs stretching over time are encoded in terms of spike rate, so are naturally washed out by the presence of noise in the network if they are not strong enough. Though we are musing outside of the realm of directly verifiable observations, the interpretation by [54] that neuronal noise facilitates a "smart" encoding scheme which enables correlated inputs to pass and causes uncorrelated inputs to decay, further illuminates the means through which neural variability could be useful in the brain.

Now that a general survey has presented noise as an integral part of understanding the function of the nervous system, the next section will offer a more focused review. To clearly define the scope of this work, note that noise in the nervous system comes from many sources [20]: chaos from coupled networks, non-stationary modulation of the nervous system, sensory input, synaptic noise, stochastic ion channel conductance, ion diffusion and many others. Here, the focus will be on how stochastic fluctuations in the macroscopic conductance of the cell membrane to ions affect the dynamics of a single neuron. The next section will outline the fundamental results in computational neuroscience which provide a foundation for this specific interest, and contextualize the review of more modern research which lies ahead.

1.2 Foundational Results of Computational Neuroscience

It is difficult to overstate the contributions of Alan Hodgkin and Andrew Huxley to the field of Electrophysiology. Their history of successful collaboration began with the first direct measurement of an action potential [30]. Following an almost decade long hiatus as a result of the Nazi invasion of Poland and the subsequent declaration of war [55], they reunited and went on to write a Nobel Prize winning series of papers which "...not only [produced] our understanding of how voltage-gated ion channels give rise to propagating action potentials, but also the very framework for studying and analyzing ion channel kinetics." [55]. Using intuition and novel fitting techniques, Hodgkin and Huxley identified the conductances which induce all of the essential currents in giant squid axion, and gave a complete dynamical description of the action potential [31]. Their insight was packaged into a system of 4 non-linear coupled differential equations which fit experimental data remarkably well, and now serves as the fundamental model of neuroscience.

Hodgkin and Huxely's insight also paved the way for other Nobel prize winning work in Electrophysiology. At the time of Hodgkin and Huxley's publication it was unclear how ions flowed through the cell membrane, and how such a passageway changed between conducting and non-conducting states in response to voltage fluctuations. The work of [44], which developed techniques for the precision measurement of individual ion channels, won both authors a Nobel Prize for experimentally verifying that current moves through channels which change stochastically between discrete conducting and non-conducting states. Though neurons were known to contain some element of randomness, this result made it clear that overcoming, mediating, or utilizing ion channel stochasticity was central to the operation of the neuron. Interestingly, the mechanism by which ion channels change between conformational conducting and

non-conducting states was not fully understood until [15] won Roderick MacKinnon the Nobel Prize in Chemistry in 2003. The paper was the first to detail the molecular structure of the proteins and explicitly identify the ion channel gates which Hodgkin and Huxley predicted with quantitative modeling almost a half century earlier.

Difficulty in visualizing Hodgkin and Huxley's four dimensional model of the giant squid axon, as well as its inherent mathematical complexity have led to a number of reduced, or lower dimensional models. Fitzhugh introduced an early reduced model [22] which began with the Van der Pol's equation for a relaxation oscillator, and modified it to produce nerve like dynamics. The model was later extended by Lecar and Nossal [37], who incorporated two-dimensional Brownian motion to derive expressions for stochastic spiking as a function of stimulus strength and duration. In both cases, the well established Hodgkin-Huxley model was discarded in favor of a lower dimensional model which would more readily yield insight, at the expense of being less experimentally grounded. Almost ten years later, a simple system was discovered in the barnacle giant muscle fiber by Morris and Lecar which was well approximated by a two gate, three dimensional model [43]. Even though the form of the Morris-Lecar equations were simple, they nonetheless reproduced most of the oscillatory voltage behavior of the barnacle muscle, and established that two persistent conductances are sufficient to create a physiologically representative neural model [43]. The central model of this project (defined in section 2.1.2) is based on the Morris-Lecar model, and utilizes its relative simplicity to rapidly develop illuminating tools so that they can be applied to more complex Hodgkin-Huxley type models in the future.

In 1979 Skaugen et al. used a stochastic implementation of the Hodgkin-Huxley equations to investigate how the spike rate, variability, and timing distribution vary as a function of sodium or potassium channel numbers [57]. They found that the stochastic model had a lower current threshold than the deterministic model and exhibited burst firing [57]. Though this study was mostly (as is much of the literature)

focused on firing rate, they do produce one figure which shows the dependence of interspike interval variability as a function of channel number. Figure 11 from [57] shows how the the standard deviation of the interspike interval (ISI) changes as a function of the number of sodium or potassium channels when the other channel type is deterministic (or is in the infinite channel number limit). A central goal of this project is to understand how spiking variability is altered by the channel configuration of a neural model (where channel configuration in this project will mean the total number of channels of two given types). To meet this end, various three dimensional versions of Figure 11 from [57] have been produced in Figures 4.4, 5.4, 5.6, and 5.7. Since the stochastic nature of each channel type may interact in unexpected ways to determine the ISI variability, and since real cells have finite numbers of channels, this project improves upon [57] by computing the ISI variability for a large range of finite channel configurations with two stochastic channel types. While the figures produced in this project originate from a different model, and are visualized in an unique way, the same characteristic "shelf" for low channel numbers which drops off suddenly as channel numbers increase is seen in our figures as well as in [57] which appeared more than 30 years ago.

1.3 Experimental and Computational Studies Which Illuminate the Role of Channel Noise in Neural Dynamics

The results of this project illuminate how the inherent stochasticity of ion channels is transformed into various levels of electrical description (conductance, current, voltage) to determine the variability of spike timing. While it is clear that an understanding of neural dynamics is central to cognition, the connection between the stochasticity of ion channels and the character of the neuron has yet to be made. In this section, we will present computational, theoretical, and experimental results which show that ion

channel stochasticity is an important determinant of the functional properties of the neuron. The demonstrated importance of stochastic modeling in the literature will motivate a general study of the various forms and magnitudes with which ion channel noise can be implemented in excitable cell models, and our approach of quantifying the consequences of these implementations.

1.3.1 Experimental

To provide context for how noise affects the functional properties of neural dynamics, we first focus on characterizing the channel noise itself. The work of [12] quantifies the contributions of Na^+ and K^+ channels in cultured hippocampal neurons to the power spectral density (PSD) of the cell's current and voltage signals. Current noise was shown to exhibit generalized $1/f$ behavior, and contributions of each channel were inferred by applying blocking agents to the cell. TTX removed contributions of Na^+ current noise and the corresponding PSD became dominated by $1/f^2$ signal. Contributions of K^+ were illuminated by replacing, intracellular K^+ with Cs^+, and also showed a significant reduction in the current PSD at depolarized voltages.

With these results in mind we may move one step upward in causality, from the inherent noise properties of channels as described by [12] to the effects of this noise on the functional properties of excitable cells. Beginning with an extreme example, [33] argues that action potentials can be generated by single channels in small cultured hippocampal neurons. Measurements of the cells yielded "plateau potentials" which began and ended with an exponential rise and decay in voltage, with time constants close to the membrane time constant. Additionally, the plateau amplitudes are consistent with currents resulting from a single high conductance channel. All of these point to single stochastic channel openings. In Figure 1B, it is clear that a small number of channels remain open as the cell slowly charges, until reaching an action potential. While this example is extreme, it highlights the

notion that the nervous system contains (and perhaps utilizes) components which are critically determined by the stochasticity of channel openings. Johansson and Arhem suggest that this cell could be used: as a random number generator (which is reminiscent of [5] from section 1.1), for maintaining a minimum level of activity to sustain neural survival, or for establishing synaptic connections in early development.

When confronted with the idea that noise critically determines the dynamic properties of excitable cells, it is natural to ask: How would the dynamics of the cell be different if stochastic channel conductances were exchanged with deterministic ones? If the behavioral states of the cell, or its processing abilities were notably altered the importance of noise would be clear. The use of dynamic clamp techniques [49, 56] and the work of [13] offer an answer to this question. In [13] the currents of entorhinal stellate neurons originating from persistent Na^+ channels were pharmacologically blocked, and dynamic clamp was used to knock-in channels using a stochastic or deterministic model. Stochastic knock-ins were the only model which produced experimentally observed peri-threshold oscillations, and had better phase reliability than their deterministic counterpart. The alteration of the cells' dynamics in the presence of a stochastic conductance model provides a direct demonstration that channel noise is an important determinant of the functional properties of entorhinal stellate neurons.

1.3.2 Computational

Results of computational studies can be roughly placed in two categories. The first offers a simulation based analysis of how stochastic modifications of the Hodgkin-Huxley model alter its dynamics, while the second implements the Hodgkin-Huxley equations in more complex or morphologically detailed models to strengthen the connection between simulation and experiment. We will begin with a review of simple implementations and then proceed to the more complex models.

After characterizing fast and slow Na^+ and K^+ currents from medial entorhinal cortex (MEC) neurons using voltage clamp methods, [65] derived a hybrid stochastic-derministic model. Their experimentally based model showed that channel noise increases the range of the MEC stellate cell's behavioral states (Fig. 2 of [65]) and that noise level critically determines which states the system can occupy. Additionally, they argued that a small number of persistent Na^+ channels could be an important determinant of cellular level dynamics.

Similarly, using a "realistically large" number of channels, [54] offers an overview of the effects channel noise has on spiking dynamics of a stochastic Hodgkin-Huxley model. They observed that channel noise in the stochastic model caused DC inputs to reduce measures of spike time reliability and precision in comparison to AC inputs. They proposed that DC currents caused the inherent stochasticity of the channels to determine the spike times whereas AC inputs "override" the channel stochasticity and are transferred more directly to spikes times. Figure 1 of [54] shows how this feature of channel noise emphasizes the distinction between DC and AC inputs in comparison to the deterministic model. In addition, [54] noticed, just like [65], that channel noise augments the behavioral states of the system to more closely match experimental observations. Exclusive features of the stochastic model used by [54] include: voltage oscillations between spikes and sub-threshold inputs, spontaneous spikes for subthreshold inputs, and "missing" spikes for suprathreshold inputs. These new behavioral states have emerged as a result of channel noise mixing the otherwise distinct states of the system (subthreshold and suprathreshold), and thereby increasing its richness.

1.3.3 Beyond Stochastic Hodgkin-Huxley Models

Extensions of the comparatively simple Hodgkin-Huxley equations into more complex models has been done with some success. The models generally offer more complexity

by either adding additional conductances to the Hodgkin-Huxley model [51], or by considering multi-compartment models [8, 19].

To create a fast and accurate simulation of cerebellar granule cells, [51] implemented a 6-conductance stochastic differential equation model with calcium dynamics. The model exhibited many features which could not be simulated in a deterministic model, and did so much faster than studies like, [19], computed using exact algorithms over a number of months. To understand how cell morphology influences the functional consequences of ion channel noise, another study performed by [8] created and tested tools to simulate multi-compartment systems where ion channels are distributed across complex morphologies. Figure 6 from [8] details how stochastic channel gating manifests in the voltage variability for many different cell types, as well as how this voltage variability propagates along the dendrite from the soma. A similar study was done by [19] with a focus on how stochastic properties of neurons alter information transmission in thin axons.

1.4 Analytical Studies of Neural Models

With the exception of the simplest neural models, analytical methods have not been particularly fruitful in illuminating neural dynamics in the literature. A great deal is known about the integrate-and-fire model [6, 7, 60], but theoretical results for Hodgkin-Huxley type models which are relevant to this project are limited. However, the work of [58] describes the voltage noise of a Hodgkin-Huxley type model by breaking approximations of Hodgkin-Huxley ion channel kinetics into two categories. Passive linear approximations replace a full kinetic scheme with a single conductance, while quasi-active linear approximations replace the kinetic scheme with a "phenomenological impedance". Both approximation types in [58] are used to derive voltage power spectral densities, and are compared with Monte Carlo simulations.

The quasi-active linear approximations was determined to be the most accurate of the two.

Another notable result which gives the mean firing rate for subthreshold inputs was derived by [9]. Building off of the work of [37], [9] uses a more modern approach by thinking of stochastic action potentials in terms of a barrier escape problem. The barrier escape perspective enables the application of tools from statistical mechanics which facilitate a derivation of the mean firing rate as a function of cell area, and estimates of the interspike interval (ISI) distribution. Unlike [37], who linearized around the threshold, [9] used a reduced model which retained non-linearity.

1.5 Stochastic Ion Channel Kinetics Schemes: Connections and Their Consequences

Hodgkin and Huxley outlined a dynamical model of the essential conductances and currents present in the squid giant axon [31]. Their model was the primary perspective by which scientist and theoreticians understood the neuron until advances in the patch clamp technique enabled the measurement of individual channels [44] and showed that channels were stochastic elements which opened and closed randomly. New observations of the microscopic currents made by [44] were then paired with pre-existing knowledge about macroscopic currents from the Hodgkin-Huxley equations to create an experimentally based stochastic model of the squid giant axon. Instead of using deterministic rate equations to propagate the average states of the sub-units through time, a new model where the open channels of each type could be viewed as a continuous time Markov chain with nonlinear, voltage dependent transition probabilities was proposed. While this new perspective faithfully represents the stochastic nature of system, it comes at the cost of independently tracking the number of channels in each state. As a result it is both computationally expensive and difficult

to analyze mathematically. Researchers who sought to understand neural dynamics through computer simulation were forced to make a difficult choice: use a fast, simple, but outdated [44] deterministic algorithm or a slow, complex, but accurate model.

In the following sections, we begin by reviewing various implementations of the Markov model in the literature. We then move to Fox and Lu's famous and controversial derivation of a stochastic differential equation (SDE) which attempted to satisfy the calling for a fast and accurate stochastic model of the Hodgkin-Huxley equations. Afterwards, the resulting conflict in the literature regarding the accuracy of the Fox and Lu's SDE is detailed, and then this section ends with a resolution that explains the confusion and outlines the correct approach, following [27].

1.5.1 Simulating Discrete Markov Models of the Hodgkin-Huxley Neuron

To review different computer implementations of the model, the conventions of [42], which provides a concise survey of past work, will be used. Each implementation belongs to one of two categories. The first category of algorithms, called channel state tracking algorithms (CST) [59, 50], are focused on *sub-units*. The probabilistic transitions happen through gates within the channel, and afterwards the states of the sub-units can be tallied up to give the state of the channel. The second category, called channel number tracking algorithms (CNT) [9], originally developed by [26], utilizes a different perspective which is focused on *channels*. The fundamental element is no longer sub-unit state, but the number of channels that have a given sub-unit state (e.g. there are 12 sodium channels with one h gate open and two m gates open). The essential consequence of this is that certain sub-unit states within a channel are no longer distinguishable. For instance, there are four indistinguishable ways that a potassium channel could have one gate open. So, transitions in this case are between channels with 1 open gate to channels with 1 ± 1 open gates.

The work of [42] compares three of the essential implementations of the model which have appeared in the literature: the CST algorithm of Strassberg et al. [59], and Rubinstein et al. [50], as well as the CNT algorithm of Chow et al. [9]. Note that [42] is not a comparison of Hodgkin-Huxley models per say, but a comparison of kinetic simulation algorithms used in stochastic Hodgkin-Huxley models. This is made clear by the fact that they perform the analysis using a sodium channel model that neglects the role of potassium. Though this places the results one step away from their application, it acts to focus the attention on comparing gating simulation schemes instead of neural models. They evaluated the models though a Monte Carlo simulation which injected a stimulus current 1000 times for various model time steps. Comparisons were then made between the number of spikes which occurred divided by the total number of stimuli, the mean value of spike occurrence times, and the standard deviation of spike occurrences times, with the conclusion being that all three models were consistent. However, the CNT algorithm of Chow et al. [9] was the most computationally efficient and robust model of the three.

It is important to note that [9] and many other sources in the literature use an approximation (referred to as the piecewise constant propensity approximation (PCPA) described in detail in section 4.2) which is largely inconsequential for systems with many channels, but many cause significant inaccuracies in smaller models. In equation 39 of [9] it is clear that the rate parameter is a constant even though the voltage is changing, whereas the analogous expression in equation 2 of [10] accounts for the changing voltage. Even though the exact algorithm has been know since at least 1983 [10], it has not been widely recognized in the neuroscience literature. The work of [1] formulates the exact representation, and uses stochastic implementations of the Morris-Lecar neuron to compare the exact and PCPA model. Figure 6 of [1] shows that the L1 distance between histograms of the exact and PCPA models becomes small after about 40 channels.

1.5.2 The Emergence of a Langevin Approximation

The following model categories have been outlined: fundamental, deterministic models which came directly from experiment [31, 43]; exact, stochastic extensions of these models [10, 1]; and PCPA of the exact, stochastic models [9, 1]. The first group does not capture the stochastic nature of a neuron, while the other two are prohibitively expensive in terms of computation as the number of channels grows. So, an unfilled niche for a fast, accurate, stochastic model whose runtime is not affected by scaling channel numbers remains.

Fox and Lu filled this niche by placing the current understanding of stochastic ion channels and the Hodgkin-Huxley model into the theoretical framework of probability theory. In doing so, they derived two stochastic differential equation models which captured the stochastic nature of neurons (under certain conditions) in a computationally efficient way [24, 23]. Additionally, their work enabled a new set of tools from nonlinear stochastic differential equations to illuminate the stochastic Hodgkin-Huxley model.

The difference between Fox and Lu's two SDEs originates from how sub-units or channels are grouped together in the derivation. The fundamental variable in the *channel-based* SDE for potassium is a vector whose kth element gives the number of potassium channels with k open subunits. For example, the 4th element of the potassium channel vector is the number of channels in the conducting state. The end result of this approach is an SDE which governs how the number of channels in each state evolves over time. Using this microscopic description of the channel states, the full dynamic equation which expresses macroscopic currents is obtained by simply substituting the new description of the channels into the Hodgkin-Huxley current balance equation. On the other hand, the *sub-unit* based model uses the fraction of open sub-units as its fundamental variable, and results in one SDE for each sub-unit. An intermediate step then multiplies the sub-units together to form the fraction of

open channels (just like the classical H-H equations), and then this derived "channel variable", represented by a multiplication of stochastic processes, is substituted into the current balance equation.

In each case the derivation follows the following format: Begin by expressing the Markov chain model of the Hodgkin-Huxley kinetic scheme in terms of a *channel-based* or *sub-unit-based* master equation. Assuming, informally, that the number of channels or sub-units is large, a system size expansion can be derived which yields a Fokker-Planck equation. The Fokker-Plank equation can then be transformed into the corresponding *channel* or *sub-unit-based* stochastic differential equation (SDE) that (when channel or sub-unit numbers are sufficiently large) yields a fast model which should faithfully represent the stochastic nature of the system. The speed advantage of the SDE approach comes from the fact that it aggregates the effects of many channel/sub-unit openings into a Gaussian stochastic process.

The primary result of Fox and Lu's work was the *channel-based* SDE, because it could be included in the current balance equation directly, and was thereby more rigorous and careful. However, there were some drawbacks. Firstly, it was not clear how this new stochastic channel based approach was related to the classical sub-unit based Hodgkin-Huxley equations (though it is now [28]). Implementation was more complex because, for instance, the channel based potassium channel SDE was 5-dimensional (one dimension for each channel state) instead of the one dimensional sub-unit approach, and the SDE contained a matrix square root which needed to be computed at each step. With this in mind Fox and Lu remark,

> ...because the deterministic equations (1), (31), and (32) are so well approximated by the Hodgkin-Huxley equations (1)-(4) it turns out to be very accurate and much simpler to implement the [sub-unit based SDE] [24].

The drawbacks of the channel-based SDE and the encouragement of Fox and Lu caused focus to shift towards the sub-unit based approach and away from the chan-

nel based approach. Slowly, over the following ten years, the research community forgot about the channel based SDE and began to discover the inadequacy of the sub-unit based approach. In the following sections we review how the research community quantitatively invalidated the sub-unit model, and then eventually came to understand fast and accurate diffusion approximations for Hodgkin-Huxley channel kinetics.

1.5.3 Conflict Over Fox and Lu's Langevin Approximations

Interested in applying the new sub-unit based SDE method which was outlined by Fox and Lu [24, 23], investigators began implementing and testing the new model against the standard Markov chain method. Subsequently several papers identified discrepancies between the SDE method which was derived by Fox and Lu and the standard Markov chain method. A discrepancy was first (to my knowledge) noted by [42]. Mino et al. outlines four variations of a 1000 channel, sodium-only Hodgkin-Huxley model, three of which are Markov chain models, and the last is Fox and Lu's SDE. After subjecting each model to monophasic and "preconditioned" monophasic current injections, they noted that the resulting ISI histograms of Fox and Lu's model differed significantly from the three other Markov chain models which were consistent [42], and that the observation persisted across many time step sizes.

This surprising result prompted a letter from [3], who criticized the implementation methods of [42]. Bruce observed that the discrete Markov models reported the number of open channel as non-negative integers whereas Fox and Lu's continuous state SDE model reports a real number, and that the implementation in [42] rounded this real number *down*. Bruce argues that the channels should be rounded to the nearest integer [3], and reproduces the work of [42] using this claim. In response to these claims [67] agreed with the observations of Bruce [3], but emphasized that though the results differed with Bruce's revision, their primary thesis remained:

Langevin-based simulation based on Fox and Lu's approximation is unreliable and should be approached with caution. They hypothesize that the approximation is unreliable since it is, "...only strictly valid for conditions near equilibrium or with linear additive noise" [67] .

Two years later a supplementary paper by Bruce [4] echoes the doubt of Fox's approximation, and further outlines the inadequacy of Fox and Lu's approximation by showing that the inaccuracies persist even when channel numbers grow. As a result of these revelations, the simulation community was left in a state of confusion. The wonder child of Fox and Lu entered into the literature with the promise of offering fast and accurate simulation methods for stochastic ion channel kinetic schemes, but became shrouded with doubt and uncertainty. So, the problem persisted; Do fast and accurate simulation schemes for ion channel kinetics exist?

1.5.4 Resolution, Effective Langevin Approximations, and Reflections on Channel Noise Models

The source of re-invigoration for the efforts of Fox and Lu came independently from their original work. To my knowledge the first effective SDE method to appear in the literature was from [40] who developed a diffusion approximation of the Markov model from first principles based on the stationary statistics of ion channels. Instead of applying noise to the sub-units, a sum of Ornstein-Uhlenbeck processes, with time and noise constants derived from stationary statistics, is added to the deterministic conductance yielding a stochastic conductance. They subjected the new SDE model to a number of tests and compared its output with the exact algorithm and Fox and Lu's SDE. Results further emphasized the shortcomings of Fox and Lu's SDE, while also establishing that an effective Langevin approximation does exist and can be readily applied. However, since the model is derived from stationary values, it becomes unreliable during the fast depolarization of the action potential or during

rapid current injections. So, though this method is clearly an improvement over Fox and Lu's sub-unit based SDE, it has some drawbacks.

Coming full circle, we arrive at the work of [27] who argues that Fox and Lu's *channel-based* SDE is the most accurate Langevin approximation. Even though Fox and Lu derived the channel-based SDE model almost two decades before [24], this paper is the first (to my knowledge) to actually implement it into a simulation and compare it to other models. They numerically show that the channel based SDE surpasses the accuracy of two types of sub-unit models and conclude that it is the most accurate. A supplementary paper by two of the same authors [28] compares the sub-unit, channel, voltage-clamp [40], and exact model, and shows that the stationary methods of [40] fall short in comparison of the Fox and Lu's Channel based SDE. A direct connection between the classical H-H model and Fox and Lu's channel based SDE is also outlined on page 4 of [40], which establishes that the SDE can be interpreted as a conductance noise model, just like the model derived from stationary statistics [40].

Though we have finally found an accurate SDE model and the primary source of inquiry has been extinguished, the fire has spread; Why did Fox and Lu's sub-unit algorithm fail? From where did the shortcomings of Fox and Lu's approach arise, and are there general principles for identifying effective approximations of stochastic ion channel kinetic schemes? A detailed derivation of the mean and variance in the fraction of open channels which results from the two approaches is given in [27]. In the channel based approach, one simply "creates" a channel by multiplying binary random variables together which represent sub-units. The channel is open if the multiplication of sub-units is one and is closed if it equals zero, thereby making the channel itself a binary random variable. To find the fraction of open channels at a given time, we simply sum over the N channels (binary random variables) and divide by N. The sub-unit approach does this in the opposite order. Namely, the number

of open sub-units of each type are summed, which gives a sum of binary random variables, and then the sum of each type is multiplied together to form the fraction of open channels. It turns out that multiplying open sub-units and then summing their product creates a different random variable than summing open sub-units and then multiplying them. Namely, the random variable created by the sub-unit approach has a lower variance [27]. Interestingly, the expectation values of these two approaches are the same, so deterministic models using each approach would be equivalent, but stochastic versions generates distinct stochastic processes. We can now look back at previous studies [42, 3, 4, 40, 28] and understand the source of the numerical discrepancy between Fox and Lu's sub-unit approach and the exact Markov chain model.

Now that we have seen what works and what doesn't, we can make some general observations. Though the source of stochasticity in ion channel conductance originates from the sub-unit, results up to this point suggest that conductance noise is more accurate than sub-unit noise. There seems to be a fundamental incompatibility between the Hodgkin-Huxley approach, which determines fractions of open channels through the multiplication of the fractions of each open sub-unit type, and the addition of noise terms in sub-unit equations. Though some have proposed corrections to the sub-unit approach [4], a successful model will likely need to alter the underlying classical model or the placement of noise terms. Thus far, the only successful Langevin approximations (to my knowledge) have used *conductance* noise [40, 24, 23, 28], which suggest that channel noise should both be implemented and thought of in these terms.

1.6 Contributions of This Project To Understanding Channel Noise and Spiking Variability

In the text above we have outlined a number of stochastic Hodgkin-Huxley models including: Clay and DeFelice's exact channel number tracking Markov chain model [10], the piecewise constant propensity approximation of this model [9, 1], as well as Fox and Lu's channel based SDE(s) [24, 23]. This project implements each of these three types of algorithms in a stochastic Morris-Lecar model to evaluate the numerical convergence of each simulation method to the others in terms of interspike interval mean, variance, and coefficient of variation (CV).

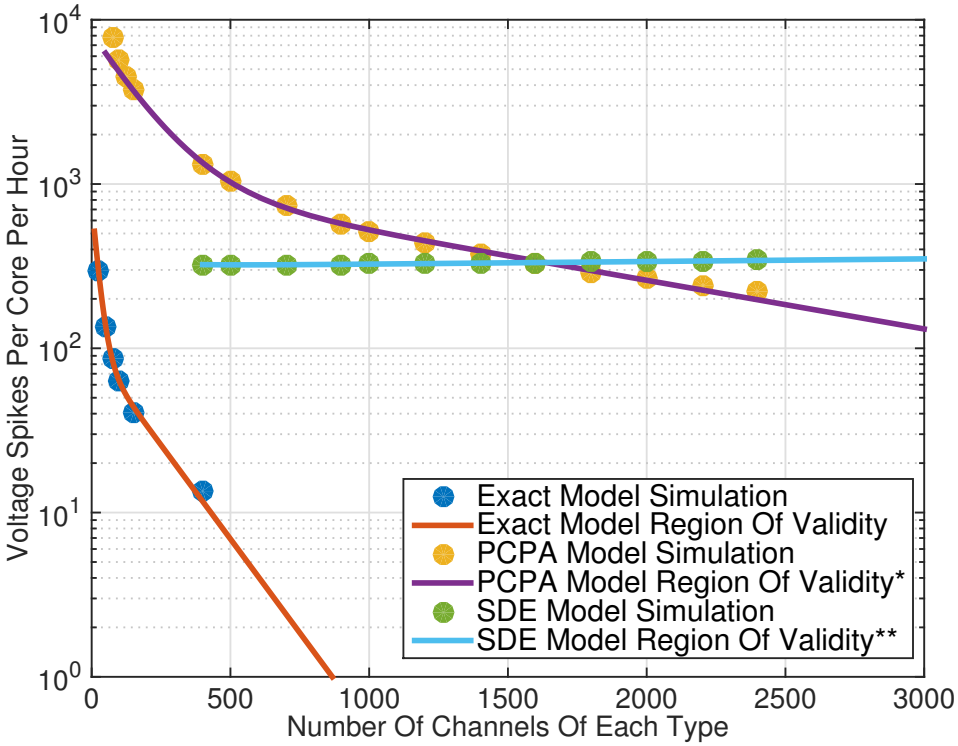


Figure 1.1: A comparison of 3 numerical implementations of the system defined in section 2.1.2 which shows how implementation runtimes and validity regions vary as a function of channel number, a central result of this project.

Though results can be depicted in three dimensions where channel numbers of each type (N_{M} and N_{N}) vary in the positive quarter of the plane (as is done in figure 5.1), for the sake of visualization, the x-axis represents the $N_{\mathrm{M}} = N_{\mathrm{N}}$ line, where the numbers of each type of channel are equal. The y-axis shows the number of voltage "spikes" each implementation yields for each CPU core per hour. Voltage spikes, defined explicitly in section 2.1.3, mark the number of system cycles which have been simulated. A central goal of this project is to understand how the statistics of the system's random period change as a function of the system's channel numbers. Simulating more cycles per unit time is desirable because it leads to more accurate statistical estimates per unit time.

The BLUE dots represent the simulation speed of an exact implementation defined in section 4.1. Its speed decreases dramatically as the number of channels are increased, which motivates the derivation of faster methods. The YELLOW dots represent the simulation speed of a piecewise constant propensity approximation (PCPA) to the exact implementation which is derived in section 4.2. An argument that the PCPA converges to the exact implementation is presented in section 4.3, but attempts at numerical validation, shown in Figure 4.4 and 4.5, suggest otherwise. Though this method is significantly faster than the exact implementation, its runtime still decreases as the channel numbers are increased. The GREEN dots represent a stochastic differential equation (SDE) model, derived in section 5.1, which approximates the exact model. Its convergence to the PCPA is validated in Figure 5.5 and shown in Figure 5.4. The advantage of the SDE model is that it has a fixed runtime (which depends only on the size of the time step) as a function of channel numbers. Each dot represents the average of at least 47,000 spikes.

The last model in this project, derived in section 6.4, and numerically validated in Figure 6.1 (with limited success), becomes a good approximation to the SDE model for channel numbers exceeding ∼100,000 of each type, and is a closed form expression which yields the ISI mean and variance, without the need for simulation, instantaneously.

The RED line represents the region where the exact model is valid. Since the exact model is valid for all channel configurations, the region is restricted only by simulation speed. * The PURPLE line represents the region of validity for the PCPA model as determined by *voltage histogram* convergence between the exact and PCPA models in [1] (which occurs after ∼50 channels of each type are used), since this project's approach to establish convergence in *ISI statistics* was not successful (see section 4.3). Note that convergence of *voltage histograms* does not necessarily imply convergence in *ISI statistics*. ** The TEAL lines represent the region of validity for the SDE model, whose convergence to the PCPA model is depicted in Figure 5.4, and shown to be statistically significant in Figure 5.5. As described in section 4.3, there were some issues establishing convergence between the exact and PCPA implementations, so note that the SDE model was validated using the unsubstantiated PCPA model. Each line was found by fitting an exponential model, $f(x) = ae^{bx} + ce^{dx}$, to the data points depicted in validity regions of figure.

Additionally, the numerical convergence of the SDE approximation to a novel analytical solution, derived by extending the results of [17], is determined. In the process of establishing these results, large scale simulation tools, as well as efficient error estimation techniques were developed.

Previous studies have shown how the system changes as channel numbers are increased, and others have shown how one model converges to another. Here, the dependence of spike rate and variability on channel numbers is shown across an entire surface, starting from the smallest number of channels where spiking is well defined and moving all they way into the limit where stochastic models converge to their deterministic counterparts. The expansiveness of this surface provides an environment to systematically analyze where each of these four aforementioned models converges to the others, and the power and efficiency of an implementation of these models on a high performance computing cluster enables the results to be presented with a scale and accuracy not previously demonstrated.

The most prominent contribution is a practical one. At present, given a cell model with a certain number of channels, it is unclear both a.) where the given number of channels places the model in terms of interspike interval statistics and b.) which of the four simulation methods is fastest for the desired accuracy. This project's description of how each model converges across the entire channel number - interspike interval surface answers both of these questions (though inconsistencies arose in the model defined in section 4.3). Additionally, for models with more than about 10^5 channels, we provide an analytical solution which can be used instead of solving an SDE.

Perhaps more importantly, the visualization of the entire surface depicts the system across its entire range of channel configurations, and thereby fully characterizes its behavior as a function of channel number. Though the Morris-Lecar equations mostly offer theoretical insight, the implementation of each version on a high performance computing cluster has paved the way for more physiologically relevant models.

With this in mind, offering the complete picture to the research community enables scientists to visualize the various features of the surface and make quick judgements about how sensitive the system is to changes in channel number. For instance, if the average number of channels of a given cell type is right between two different features of the ISI surface, then cell to cell variability could significantly alter the dynamics of that type of cell, whereas the dynamics of a cell whose average number is placed firmly in one region would remain the same.

Chapter 2

Overview of Central Model and Motivations

Computation in the brain results from networks of neurons which propagate information by sending sequences of electrical discharges called action potentials. An action potential, or voltage spike, is caused by a flux of certain ions across the cell membrane of a neuron. Ion channels, the primary pathway for ion movement across biological membranes, determine the effective conductance of the membrane, and can be modeled as discrete state stochastic elements which open and close in a random, voltage-dependent way to enable the passage of current through the cell. These channels, which determine the action potential profile of each cell, populate the membrane of a neuron in many different types, all of which are coupled by a common membrane voltage. Each type of channel conducts an associated ion and has a voltage-dependent open probability. The number, type, and voltage dependence of ion channels in a neuron determine the conditions under which the cell transmits signals from its inputs.

This project will investigate the dynamics of a modern, Markov process version of Morris-Lecar's neural model, originally developed to describe voltage oscillations in the barnacle giant muscle fiber [43]. The presentation will begin by describing a

simple model of stochastic currents and then using the simple model to construct the exact hybrid Markov model of the Morris-Lecar neuron (which will be progressively approximated in chapters 4, 5, and 6). Next, the solutions of the exact hybrid Markov model of the Morris-Lecar neuron will be compared to experimental data, and methods to quantify the characteristic "spiking" which they both exhibit will be introduced. Lastly, the central questions of this project will be presented, and the methods to answer them will be outlined.

2.1 Definition and Dynamics of the Exact Hybrid Markov Model of the Morris-Lecar Neuron

2.1.1 A Simple Model for Stochastic Currents

In the Markov chain description of the Morris-Lecar model, a single channel, $c(t)$, is a binary stochastic process which moves between conducting ($c(t) = 1$) and non-conducting ($c(t) = 0$) states with probabilities related to the membrane voltage, V. When the channel is in a conducting state, a unitary current per unit area, i, flows through the channel, and is defined by the specific unitary conductance of the channel g, the membrane voltage V, and the reversal potential V_{R},

$$i(t) = c(t)g(V(t) - V_{\mathrm{R}}). \tag{2.1}$$

Note that parameters are given in units per area to make scaling the system easier in the future. The total current per unit area, I, through a population of N channels is simply a sum over the specific unitary currents of its members,

$$I(t) = g\,(V(t) - V_{\mathrm{R}}) \sum_{k=1}^{N} c_k(t). \tag{2.2}$$

To simplify discussion we can define a new discrete state stochastic variable which counts the number of open channels as a function of time,

$$n(t) = \sum_{k=1}^{N} c_k(t), \tag{2.3}$$

so that the total current density through a population of channels is,

$$I(t) = n(t)g\left(V(t) - V_{\mathrm{R}}\right). \tag{2.4}$$

Additionally, we can consider the total specific conductance $G = Ng$, and express the current densities in terms of the fraction of open channels,

$$I(t) = \frac{n(t)}{N} G\left(V(t) - V_{\mathrm{R}}\right). \tag{2.5}$$

2.1.2 The Exact Hybrid Markov Model of the Morris-Lecar Neuron

The Morris-Lecar model assumes that three distinct current densities are present,

$$I_{\mathrm{soma}} = \frac{n_{\mathrm{M}}(t)}{N_{\mathrm{M}}} G_{\mathrm{M}}(V - V_{\mathrm{M}}) + \frac{n_{\mathrm{N}}(t)}{N_{\mathrm{N}}} G_{\mathrm{N}}(V - V_{\mathrm{N}}) + G_{\mathrm{L}}(V - V_{\mathrm{L}}), \tag{2.6}$$

where each of the three groups of terms above represents a distinct current per unit area. M-type channels pass an inward calcium current, N-type channels pass an outward potassium current, and L-type channels pass a nonspecific leak current. In each case, G represents the maximal specific conductance of the membrane to each type of current, and, for $k = [M, N, L]$, V_k is the reversal potential for the kth current. For M and N type currents, N_k is the total number of channels, and $n_k(t)$ is a discrete stochastic process which gives the number of those channels which are open as a function of time. More specifically, $n_k(t)$ is defined to be the difference of

two non-homogeneous Poisson processes, with mean functions $m^k_{\text{open}}(t)$ and $m^k_{\text{close}}(t)$,

$$n_k(t) = N^k_{\text{open}}(\, t,\, m^k_{\text{open}}(t)\,) - N^k_{\text{close}}(\, t,\, m^k_{\text{close}}(t)\,). \tag{2.7}$$

N^k_{open} counts the total number of opening events for a channel of type k after t time units, and N^k_{close} counts the closing events. Their difference gives the total number of open channels as a function of t. For ease of notation we will henceforth suppress the channel type sub/superscript k when it is not essential. Each channel type has a unique pair of experimentally measured, voltage dependent functions which determine the opening,

$$\begin{aligned} \alpha_{\text{M}}(V) &= M_\infty(V)\lambda_{\text{M}}(V), \\ \alpha_{\text{N}}(V) &= N_\infty(V)\lambda_{\text{N}}(V), \end{aligned} \tag{2.8}$$

and closing,

$$\begin{aligned} \beta_{\text{M}}(V) &= (1 - M_\infty(V))\lambda_{\text{M}}(V), \\ \beta_{\text{N}}(V) &= (1 - N_\infty(V))\lambda_{\text{N}}(V), \end{aligned} \tag{2.9}$$

rates of the channels (as shown in Figure 2.1). The parameters[1][2][3] that define the component functions above, which are described fully in [43], are standard Morris-Lecar parameters that (along with the other parameters defined in the Nomenclature section) give a globally attracting limit cycle, or a fixed point near a Hopf bifurcation. These rate functions, together with the number of channels available for transition, define intensity parameters, λ (not to be confused with rate constants λ_{M} and λ_{N}), whose integral defines the mean functions,

$$m_{\text{open}}(t) = \int_0^t (\, N - n(s)\,)\, \alpha(\, V(s)\,) ds = \int_0^t \lambda_{\text{open}}(s) ds, \tag{2.10}$$

$$m_{\text{close}}(t) = \int_0^t n(s)\beta(\, V(s)\,) ds = \int_0^t \lambda_{\text{close}}(s) ds, \tag{2.11}$$

[1] $M_\infty(V) = (1/2)(1 + \tanh((V - V_1)/V_2))$, $\lambda_{\text{M}}(V) = \bar{\lambda}_{\text{M}} \cosh((V - V_1)/2V_2)$
[2] $N_\infty(V) = (1/2)(1 + \tanh((V - V_3)/V_4))$, $\lambda_{\text{N}}(V) = \bar{\lambda}_{\text{N}} \cosh((V - V_3)/2V_4)$
[3] $V_1 = -1.2$, $V_2 = 18$, $V_3 = 2$, $V_4 = 30$, $\bar{\lambda}_{\text{M}} = 0.4$, $\bar{\lambda}_{\text{N}} = 0.04$

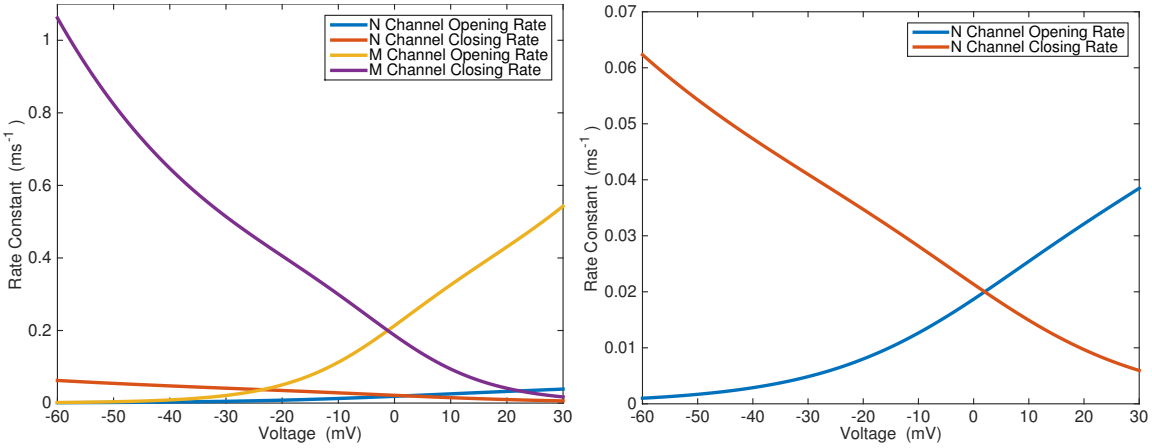


Figure 2.1: The rate functions defined in equations 2.8 and 2.9 (where the component functions are defined by $V_1 = -1.2$, $V_2 = 18$, $V_3 = 2$, and $V_4 = 30$ in [43]) are graphed across a range of voltages the system occupies. (right) shows all four rate functions and facilitates a comparison of magnitudes (left) shows a magnified graph of the N channel rate functions. Note that M channels have a positive reversal potential $V_{\mathrm{M}} = 120$ mV, N channels have a negative reversal potential $V_{\mathrm{N}} = -60$ mV, the rate functions for the M channel are approximately 10 times larger than those of the N channel (so they react to voltage changes much faster), and that there is a small voltage difference between when M and N channels start opening or closing. These four factors combine to create the periodic dynamics seen in Figure 2.4, and are an example of a very general phenomenon in neuroscience. The biophysical details of why the periodic behavior arises can be found in [48].

where N is the total number of channels of a given type.

Equation 2.6 shows that when voltage is fixed the stochasticity of open channels is passed simply to the somatic current through an affine transformation. An expression of the voltage (the real determinate of cell dynamics) can be found by integrating the current, or through the following differential equation,

$$\begin{aligned} C\frac{dV}{dt} &= I_{\mathrm{app}} - \frac{n_{\mathrm{M}}(t)}{N_{\mathrm{M}}} G_{\mathrm{M}}(V - V_{\mathrm{M}}) \\ &\qquad - \frac{n_{\mathrm{N}}(t)}{N_{\mathrm{N}}} G_{\mathrm{N}}(V - V_{\mathrm{N}}) - G_{\mathrm{L}}(V - V_{\mathrm{L}}) \\ &= I_{\mathrm{app}} - I_{\mathrm{soma}} \end{aligned} \tag{2.12}$$

where C is the specific capacitance of the membrane, and I_{app} is an applied current. By noting that the left hand side represents the total current which passes through the membrane of the cell, and that the right hand side is the sum of individual currents which make up the total current, the differential equation above can be understood as a current conservation equation.

If channel transitions occur at random times $\{t_1, ..., t_{\mathrm{N}}\}$, then V(t) is a determin-

istic, continuous function within each interval between t_k and t_{k+1} for each k (here k denotes the index of successive times, not the channel type). In fact, since the differential equation for V is linear, the solution between t_k and t_{k+1} is exponential, where the exponential rate is determined by the open channel numbers, and the entire solution is piecewise exponential. Current then becomes a hybrid Markov process which mixes the discrete channel state process, given by the difference of two non-homogeneous Poisson process, with a continuous, conditionally deterministic process, the voltage. Figure 2.2 illustrates the interdependence of current, voltage, and channel state.

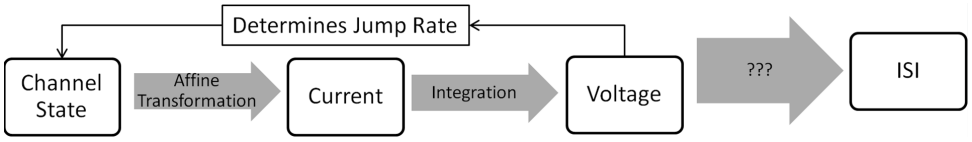


Figure 2.2: Causal structure of neural dynamics. Channel state and voltage together determine the instantaneous ionic current through an affine transformation defined by equation 2.6. The resulting current induces a change in voltage through charge integration defined by equation 2.12. Since the transition rates of the channels react to the voltage change through equations 2.10 and 2.11, a causal loop is formed. Though it is clear in the model defined in section 2 that the noise begins at the level of the channel, the presence of this feedback loop obfuscates the relationship between observations of variability in channel state and voltage. Furthermore, the causal loop propagates the voltage in an undetermined way to induce a point process of "spikes" or "interspike intervals" (ISIs) (defined in equation 2.13) which play an important physiological role. The way in which the inherent stochasticity of the channel state moves through the causal loop to determine the voltage, and thereby the ISI distribution, has important biological implications, and is a central concern of this project.

2.1.3 Dynamics, Spiking, and the Interspike Interval Distribution

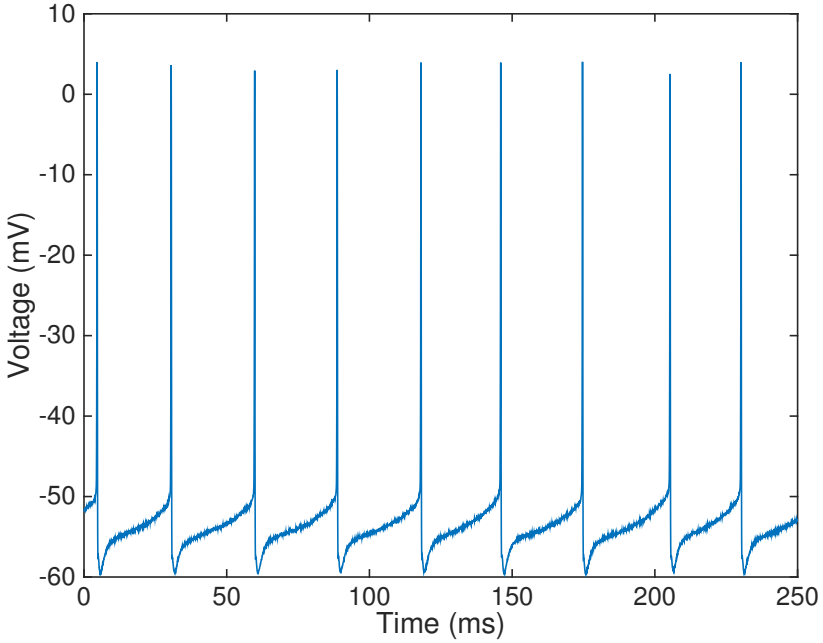


Figure 2.3: Measurement of voltage spiking activity from mouse Purkinje cells in acute cerebellar slices. Courtesy of Prof. David Friel, Department of Neurosciences, CWRU School of Medicine. For this particular recording of 791 spikes the firing rate was 38 Hz, with mean ISI of 25 ms, and ISI standard deviation of 2.8. Note the variability of the interval durations between successive spikes, and the small fluctuations visible in the voltage traces during the gradually rising phase between spikes.

Once a neuron reaches a threshold potential, it begins to rapidly increase in voltage to a maximum called a spike, firing, or action potential. After reaching a maximum, the system rapidly decreases in voltage, and resets back to some starting potential. When the stochastic Morris-Lecar neuron is driven by an applied current, the ion channel currents respond by producing a regular (although not quite periodic) train of action potentials. Our study is motivated by Purkinje cells which also fire approximately periodically, and are thought to act as pacemakers within the cerebellar motor control circuit. Such a pacemaker may be used to control oscillatory processes, or amplify and modulate passing signals in the nervous system.

We can quantify the regularity of approximately periodic spiking by considering the distribution in times between adjacent spikes, called the interspike interval (ISI)

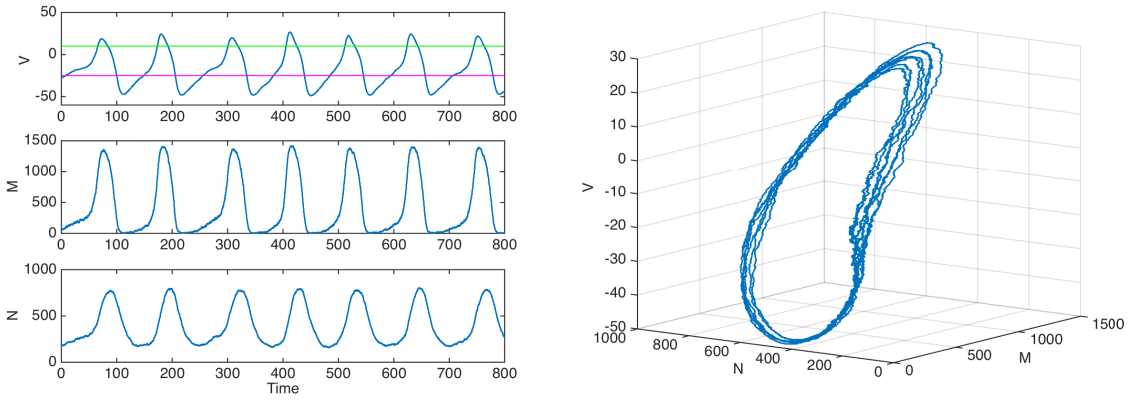


Figure 2.4: The system defined in section 2.1.2 is approximated and simulated (with an Euler-Maruyama method using $N_{\mathrm{N}} = 1500$, $N_{\mathrm{M}} = 1500$, and $\Delta t = 0.0043$) using methods developed in chapter 5. The plot shows the combined dynamics of voltage, the number of open M channels, and the number of open N channels. The system gives a stable limit cycle which is perturbed by relatively weak noise, so that very small perturbations of the trace accumulate to noticeably alter the path of the system through state space. There are a few notable differences between these voltage traces and the experimental measurement of Purkinje cells shown in Figure 2.3. Firstly, the parameters given in section 2.1.2 place the system near a Hopf bifurcation and give them a different "form" which is probably distinct from the characteristics of Purkinje cells [21]. Secondly, note that the model's entire voltage trace is uniformly about 15 mV more than that of Figure 2.3. The shift results from the model's dynamics being based on Ca^{2+} which has a higher reversal potential, $V_{\mathrm{M}} = 120$ mV, than ions which dominate the cells in Figure 2.3. Note that the high number of channels used in this model $N_{\mathrm{N}} = 1500$, $N_{\mathrm{M}} = 1500$ induces a regularity in the voltage traces, so that the two-threshold spike finding method discussed later in this section is no longer needed. Thresholds are indicated with a green line at $V_{\mathrm{Th}_1} = 10$ mV and a magenta line at $V_{\mathrm{Th}_2} = -25$ mV at the top of the left figure.

distribution. More explicitly, a repetitively spiking system yields a point process of spike times $\{t_1, ..., t_N\}$, which can be re-expressed as $N - 1$ ISIs where,

$$\mathrm{ISI}_k = t_{k+1} - t_k, \quad \text{for } 1 \leq k \leq N - 1. \tag{2.13}$$

Viewed as $N - 1$ realizations of a random variable, we can consider the mean and variance of the ISI distribution, and thereby quantify a cell's spiking regularity.

The central role of the ISI distribution in this project necessitates the development of a numerical method for identifying spikes, determining the times of each spike, and then calculating the associated interspike interval distribution. One simple method to find spikes is to pick a voltage threshold, V_{Th}, and then search the time series for pairs of points such that one point is below the threshold, and a consecutive point is above the threshold. The crossing time can then be found to first order in time step Δt by linearly interpolating the crossing between the two consecutive points.

This method was used throughout the project (with $V_{\mathrm{Th}} = 10$ mV, and sometimes without linear interpretation) in cases were the analysis involved a small number of ISI samples (like in Figure 2.6).

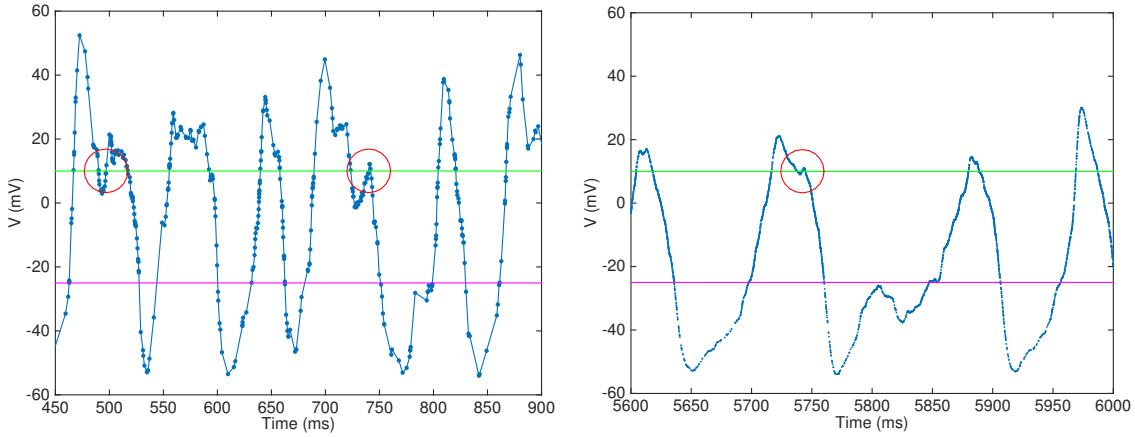


Figure 2.5: Trajectories of the system defined in section 2.1.2, simulated with (left, $N_{\mathrm{N}} = 10$ and $N_{\mathrm{M}} = 10$) and (right, $N_{\mathrm{N}} = 100$ and $N_{\mathrm{M}} = 100$), using the methods outlined in section (left, 4.1.3) and (right, 4.2). The voltage path on the right was artificially shift downwards 3.5 mV to clarify this example. Since the algorithms only record data when the state changes, each point represents a channel transition. Note that the density of points, and therefore the density of channel transitions is much higher on the right which was simulated with 10 times more channels (this is essential to the derivation of approximation methods in section 4.2). The thresholds discussed in this section are indicated with a green line at $V_{\mathrm{Th}_1} = 10$ mV and a magenta line at $V_{\mathrm{Th}_2} = -25$ mV. (left) is an extreme example which shows that a two-threshold spike finding method is essential for models with 10s of channels. The first and second red circles highlight cases where false positives resulting from a one threshold method can be removed by adding a second lower threshold which must be crossed between consecutive V_{Th_1} crossings. (right) As channel numbers increase the regularity in the paths increases and these case appear progressively less frequently. Still, the red circle highlights a potential false positive that would result from a one-threshold method, and further motivates the use of the two threshold method for large scale simulations.

However, a large scale simulation effort requires a much more robust method. As you can see in Figure 2.5, the stochastic nature of the system occasionally creates voltage fluctuations which would cause false positives using the simple method (though the parts or forms that constitute a "spike" are grey and somewhat subjective). Since there were a few hundred million ISIs simulated over a wide range of parameters, a more robust, two threshold method was used. The two threshold method is the same as the one threshold method, except that two spikes (as measured by the one threshold method) are only recorded in the two threshold method if they are separated by a $V_{\mathrm{Th}_2} = -25$ mV crossing. The addition of the lower threshold does not alter the ISI distribution very much after a few hundred channels, as shown in Figure 2.4, but is essential for anything below 100 total channels, as indicated in Figure 2.5.

2.2 Motivating Questions

The structure of the stochastic model outlined above motivates a question of general interest: *How does the inherent stochasticity of channels move through each level of description (Figure 2.2) to determine the variance of the ISI distribution?* Or, in mathematical terms, how is stochasticity of the underlying Poisson process transformed by equation 2.12 to affect the induced point process of spike times? Interestingly, the answer to this mathematical question may hold implications for a specific neurological disorder called episodic ataxia type 2.

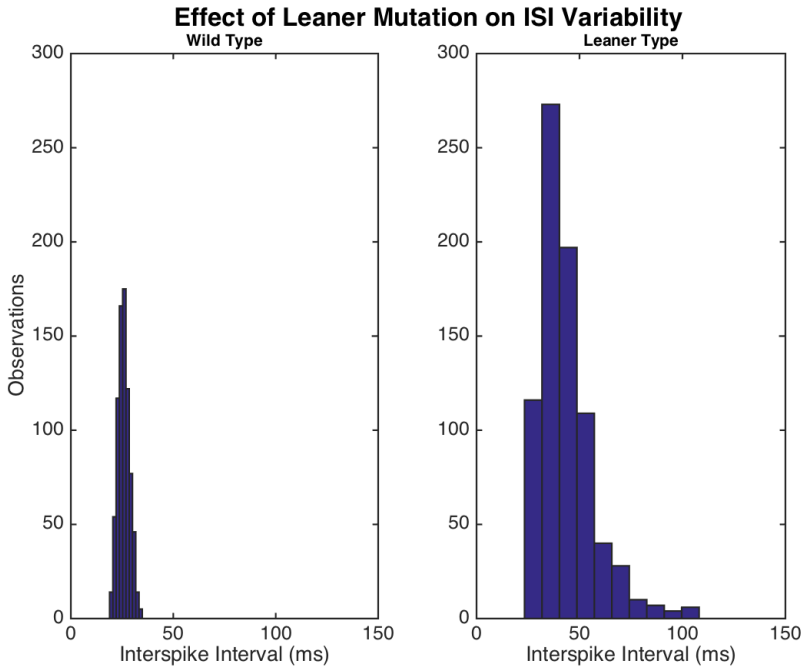


Figure 2.6: Comparison between interspike interval distributions (found using a single threshold method with $V_{\text{Th}} = -10$) of voltage spikes from wild (left) and mutant type (right) mouse Purkinje cells in acute cerebellar slices. Courtesy of Prof. David Friel, Department of Neurosciences, CWRU School of Medicine. A significant reduction in spiking regularity is measured in the mutant type cell. Figure represents 790 ISIs from each type of cell with 10 equally spaced bins starting and ending at the shortest and longest ISI respectively. A subset of the wild type voltage trace used to make this figure is shown in Figure 2.3.

Figure 2.6 shows the effect of a mutation which reduces the open probability of certain channels in dissociated Purkinje cells of mice. The mutation's effect on the electrical properties of the cell are varied and complex [46], but the interest of this project focuses on how an alteration at the level of the channel, like this mutation,

would propagate through the causal chain in Figure 2.2 to significantly broaden the interspike interval distribution. Experimental studies have established the importance of linking channel level mutations to spiking regularity, by arguing that they are the root cause of specific types of cerebellar dysfunction [63]. Thus, the motivation for this project is threefold. We are interested in: (1) how the underlying Markov process which controls channel state is transformed through each level of electrical description to alter the ISI distribution (2) How the answer to (1) illuminates the relationship between the stochasticity of each channel type and neural information processing, and (3) How the answer to (1) can enable us to get a detailed understanding of how the effects of ion channel mutations permeate through the nervous system to cause macroscopic dysfunction.

2.3 Methods

A natural way to understand how a system depends on its parameters is to systematically change the parameters and record how the system responds. Adopting this approach comes with two questions; which parameters, subject to variation, would provide the greatest insight into the dynamics of the system, and to what extent can the resulting behavior be strictly associated with variation of these parameters? Since the number of channels are the only discrete quantity in the model, and they are a critical determinate of the stochasticity of the system, a systematic evaluation of how the number of channels determines variability in the ISI distribution is both natural and useful. Furthermore, while the variation of many other parameters affects the system through necessary alterations of other variables, certain types of system scaling preserve the structure and constants of the current equation, but significantly alter the underlying Markov chain (discussed in section 3.2.1). Under these scalings, the development of ISI variability in the system can be directly attributed to changes

in the stochastic channel kinetics. A systematic evaluation of how channel numbers determine ISI variability is both illuminating and possible on these grounds, as we detail below.

A simple, brute force numerical application of equation 2.12 would only carry this line of inquiry a short distance. As the number of channels used in the system increases past a few hundred, numerical simulation becomes unfeasible. Therefore, the more general motivations of this project outlined above necessitate a tangential line of inquiry; when and how can the computationally expensive, exact Markov chain model be approximated by faster stochastic methods? To overcome the computational challenges inherent to this stochastic system, three approximate models valid for progressively larger numbers of channels are derived (which are similar to those discussed in section 1.5 and more recently [45]). By systematically simulating all four expressions of the stochastic Morris-Lecar model with various channel configurations (with limited success in the case of section 4.3), this project investigates the convergence of each approximation to the others while also furthering the general motivations outlined above. Moreover, a determination of channel configurations where the exact Markov chain description of ion channel kinetics and other conceptually different expressions of channel stochasticity yield the same ISI variability illuminates where the collective action of channels manifest in a fundamentally different way.

Verifying the numerical convergence between different stochastic models in terms of ISI variance or CV requires an expression for the standard error of the variance and CV. Once again, the brute force method for finding the standard error becomes computationally unfeasible, and another tangential challenge is presented; how can we estimate the standard error of ISI mean, variance, and CV in a computational efficient way? To overcome the computational challenges, a standard error bar approximation, valid when the number of samples exceeds 30,000, is both derived (section 7.2) and numerically verified with ISI data (section 7.3). The efficient error bar approx-

imation is paired with a powerful, parallel implementation of each model on a high performance computing cluster. Together all of these tools make it possible to numerically obtain the ISI mean, variance, and CV across many channel configurations, and several different approximations of the model with high accuracy.

While these methods stand alone in their illumination of the Morris-Lecar neuron, they also set the groundwork for a project of considerable complexity across the fields of neuroscience and mathematics. The high performance computer implementation, ISI statistic error bar approximation, large channel number closed-form solution, and channel kinetic approximation schemes are all readily generalizable to more complex, biologically relevant Hodgkin-Huxley type models.

Chapter 3

Investigating The Channel Number ISI Surface

3.1 Channel Configuration Space and the ISI Surface

Given N_{M} and N_{N}, eq 2.12, along with the associated equations and parameters presented in section 2, define a model which, when simulated over time to generate spikes, yields an associated ISI distribution, $ISI_{(N_{\mathrm{N}},N_{\mathrm{M}})}$. If S is a statistical operator (expectation value, variance, CV) then applying S to $ISI_{(N_{\mathrm{N}},N_{\mathrm{M}})}$ yields a real number. We can think of the application of S to $ISI_{(N_{\mathrm{N}},N_{\mathrm{M}})}$, an ISI distribution parameterized by channel configuration, as a function, T', which maps the channel configuration to a statistical measure of the associated ISI distribution,

$$T'(N_{\mathrm{M}}, N_{\mathrm{N}}, S) = S[ISI_{(N_{\mathrm{M}},N_{\mathrm{N}})}]. \tag{3.1}$$

When S is a statistical measure like variance, the surface defined by, $T'(N_{\mathrm{M}}, N_{\mathrm{N}}, \mathrm{Var})$, as a function of N_{N} and N_{M} will be referred to as the channel number ISI variance

surface, or simply as the ISI variance surface. The features of T' characterize how the ISI distribution of the Morris-Lecar neuron responds to various levels of noise (or different channel configurations) in terms of mean, variance, or CV. Two different stochastic expressions of the Morris-Lecar neuron (Markov chain, stochastic differential equation (SDE), etc.) will be considered numerically equivalent in a region where the T' functions associated to each model are statistically indistinguishable.

In circumstances where we would like to speak naturally about the convergence of a stochastic model, in the large channel number limit, to a deterministic one, a change of coordinates is used which places the infinite channel number limit at zero instead of infinity,

$$\epsilon_{\mathrm{M}} = \frac{1}{\sqrt{N_{\mathrm{M}}}}, \qquad \epsilon_{\mathrm{N}} = \frac{1}{\sqrt{N_{\mathrm{N}}}}. \tag{3.2}$$

The square root is present because the noise around the mean-field equations in the Gaussian approximation diminishes like $O(1/\sqrt{N})$. As a result, the closed form, infinite channel limit solution derived from a Gaussian approximation also has $O(1/\sqrt{N})$ terms. An analogous version of T', denoted T, can be defined from this perspective,

$$T(\epsilon_{\mathrm{M}}, \epsilon_{\mathrm{N}}, S) = T'(N_{\mathrm{M}}, N_{\mathrm{N}}, S), \tag{3.3}$$

so that,

$$\lim_{(N_{\mathrm{K}}, N_{\mathrm{Na}}) \to \infty} T(\epsilon_K, \epsilon_{\mathrm{Na}}, S) = T(0, 0, S). \tag{3.4}$$

In the case where S is equal to the variance operator, Var, intuitively,

$$T(0, 0, \mathrm{Var}) = 0. \tag{3.5}$$

3.1.1 Overview of Stochastic Approximation Schemes Across Channel Configuration Space

The principal goal of this project is to characterize the surface T' for a wide range of channel numbers. Using the exact Markov chain method, the system can be simulated with high accuracy for any channel configuration. However, large scale simulations of the exact Markov chain method become prohibitively slow (using MATLAB) when channel numbers of each type exceed 300 (as seen in Figure 4.2). Fortunately, the transition rates related to equations 2.10 and 2.11 can be approximated when channel numbers are as small as 100 by using a piecewise constant propensity approximation (PCPA). Instead of numerically solving an integral with a continuously varying intensity parameter λ, the PCPA sets λ to a constant, equal to its value at the time of the last transition. Still, the runtime of the PCPA algorithm is proportional to the number of channels [57], and becomes prohibitively slow after about 2000 channels of each type are used (as seen in Figure 5.1).

Fortunately, for channel numbers exceeding roughly 1600 channels the stochastic effects of each individual channel can be accurately approximated as a stochastic differential equation (SDE) driven by a pair of Gaussian process, one for each channel type. At every time step the stochastic differential equation expression of the model simply adds a vector of Gaussian random variables, scaled relative to the state of the system, to the deterministic differential equations which control the channel kinetics. As a result, when the number of each channel type exceeds about 1600 channels, every configuration can be simulated with great precision and speed. Furthermore, in the small noise regime, where channel numbers exceed 30,000, the effects of channel noise can be linearized around the limit cycle solution of the deterministic Morris-Lecar equations to yield an expression for the ISI variance in terms of integrals of functions related to the phase response curves of the mean field deterministic limit cycle, as described in chapter 6. In section 6.4.2 we investigate the relationship between the

analytic expression and numerical estimates of the coefficient of variation in the small noise limit, and find them to be in good agreement.

3.2 Scaling Across the ISI Surface

While it is now natural to ask: *How does the model respond to variations in channel number?* Care should be taken to first answer a more fundamental question: *What does it mean to vary the number of channels, and what are the physical/ mathematical implications?* In the following section we: provide a rigorous framework for thinking about scaling, illuminate what it means to scale the model, and highlight the origin of altered dynamics from scaling by identifying which parts are invariant to scaling and which are not.

3.2.1 Intensive vs Extensive Expressions of the Current Equation

Thus far, we have used a hybrid expression of the Morris-Lecar model which mixes intensive quantities like specific conductance G_{M} and extensive quantities like the total number of channels N_{M} (which grow as the area increases). While the hybrid expression simplifies both plotting and communication of the model, it obscures some of the nuance involved in scaling, so we will momentarily focus our attention on a purely intensive expression. To begin we first show that the Morris-Lecar (and by extension Hodgkin-Huxely) current equation can be expressed intensively in units of current per area. Using $\tilde{C}$ to represent the total capacitance (in microFarads) $\tilde{I}_{\mathrm{app}}$ to represent the total injected current (in nanoamperes), and $\tilde{G}_k$ to represent the total conductance through channel type k, we may write the original extensive model in

units of current as,

$$\begin{aligned}\tilde{C}\frac{dV}{dt} = \tilde{I}_{\text{app}} &- \frac{n_{\text{M}}(t)}{N_{\text{M}}}\tilde{G}_{\text{M}}(V - V_{\text{M}}) \\ &- \frac{n_{\text{N}}(t)}{N_{\text{N}}}\tilde{G}_{\text{N}}(V - V_{\text{N}}) - \tilde{G}_{\text{L}}(V - V_{\text{L}}),\end{aligned} \tag{3.6}$$

we can divide by a fixed cell area (or fixed membrane patch area) $A_F = 1\mu m^2$ to yield our hybrid expression in units of current density,

$$\begin{aligned}\frac{\tilde{C}}{A_F}\frac{dV}{dt} &= \frac{\tilde{I}_{\text{app}}}{A_F} - \frac{n_{\text{M}}(t)}{N_{\text{M}}}\frac{\tilde{G}_{\text{M}}}{A_F}(V - V_{\text{M}}) \\ &\quad - \frac{n_{\text{N}}(t)}{N_{\text{N}}}\frac{\tilde{G}_{\text{N}}}{A_F}(V - V_{\text{N}}) - \frac{\tilde{G}_{\text{L}}}{A_F}(V - V_{\text{L}}) \\ &= I_{\text{app}} - \frac{n_{\text{M}}(t)}{N_{\text{M}}}G_{\text{M}}(V - V_{\text{M}}) \\ &\quad - \frac{n_{\text{N}}(t)}{N_{\text{N}}}G_{\text{N}}(V - V_{\text{N}}) - G_{\text{L}}(V - V_{\text{L}}).\end{aligned} \tag{3.7}$$

Noting that channel number is simply the variable cell area, A, times the channel density gives,

$$\begin{aligned}C\frac{dV}{dt} &= I_{\text{app}} - \frac{\eta_{\text{M}}(t)A}{\rho_{\text{M}}A}G_{\text{M}}(V - V_{\text{M}}) \\ &\quad - \frac{\eta_{\text{N}}(t)A}{\rho_{\text{N}}A}G_{\text{N}}(V - V_{\text{N}}) - G_{\text{L}}(V - V_{\text{L}}) \\ &= I_{\text{app}} - \frac{\eta_{\text{M}}(t)}{\rho_{\text{M}}}G_{\text{M}}(V - V_{\text{M}}) \\ &\quad - \frac{\eta_{\text{N}}(t)}{\rho_{\text{N}}}G_{\text{N}}(V - V_{\text{N}}) - G_{\text{L}}(V - V_{\text{L}}),\end{aligned} \tag{3.8}$$

where ρ_k is the total density of channels of type k, and $\eta_k(t)$ is the density of open channels.

Since scaling with respect to area is a central concern of this study, the parameters in the intensive expression (C, G, I_{app}, ρ) will be taken as constants which define a class of cells that differ in only in area. While C, G, and I_{app} are always the same value, ρ_{M} and ρ_{N} can be changed to explore a different class of cells (discussed in

section 3.2.3). Given an intensive expression of the Morris-Lecar model and a cell area, A, the parameters in the associated extensive model can always be found through a simple multiplication,

$$\tilde{C} = AC, \tag{3.9}$$

$$\tilde{G} = AG, \tag{3.10}$$

$$n = A\eta, \tag{3.11}$$

$$N = A\rho, \tag{3.12}$$

and

$$\tilde{I}_{\text{app}} = AI_{\text{app}}. \tag{3.13}$$

Since all of the extensive parameters scale in proportion to A, they cancel out in the extensive differential equation, and the current equation ends up looking the same. Note that while the capacitance and the maximal conductance scale proportionally with area, the unitary conductance $\tilde{g}$ remains constant,

$$\tilde{g} = A\frac{G}{N} = A\frac{G}{\rho A} = \frac{G}{\rho}. \tag{3.14}$$

By showing that the extensive model implies the intensive model, we have also shown that the solutions, $V(t)$, from each model are equivalent, and therefore shown that if the open channel densities, $\eta(t)$, are fully determined, the voltage dynamics of the Morris-Lecar neuron (and by extension many other model neurons) are invariant under area scaling. More specifically, given that a certain density of channels are open, the way in which current (per unit area) flows through those channels to determine the voltage does not change. What does change as area scales are the statistical properties of $\eta(t)$, which governs the density of open channels through which the current can move. Mathematically, the mean-field current equation is invariant to area scaling,

but the perturbations which act upon the mean field equation are not. Here we can highlight an important point: The development of the stochastic nature of the Morris-Lecar neuron as it scales in size is completely determined by changes in the underlying Markov process which governs the channel kinetics. The only meaningful effect of increasing area is that it increases the number of stochastic elements and thereby brings the Markov process which determines the density of open channels closer to a deterministic process. So, though this study is inherently focused on experimentally observable quantities like voltage and current, changes in the stochastic properties of these quantities as the area increases is primarily a reflection of how the hidden statistics of the open channel density evolve.

3.2.2 Stochastic Ion Channel Kinetics as Chemical Reaction Networks: Convergence of Markov Chain Models to Deterministic Rate Equations

In order to understand how ion channel kinetics evolve as the number of channels increase, consider an analogous relationship with continuous time Markov chain models of chemical reaction networks. Channel state in our simple Morris-Lecar model can be thought of in terms of four distinct "reactants",

$$n(t) = \begin{bmatrix} n_{\mathrm{M_O}} \\ n_{\mathrm{M_C}} \\ n_{\mathrm{N_O}} \\ n_{\mathrm{N_C}} \end{bmatrix}, \tag{3.15}$$

which undergo "reactions"

$$\begin{aligned} M_O &\to M_C \\ M_C &\to M_O \\ N_O &\to N_C \\ N_C &\to N_O \end{aligned} \tag{3.16}$$

governed by four independent counting processes,

$$\begin{aligned} &R_{\mathrm{M_O}\to\mathrm{M_C}}\left(t, \int_0^t n_{\mathrm{M_O}}(s)\beta_{\mathrm{M}}(V(s)))ds\right) \\ &R_{\mathrm{M_C}\to\mathrm{M_O}}\left(t, \int_0^t n_{\mathrm{M_C}}(s)\alpha_{\mathrm{M}}(V(s)))ds\right) \\ &R_{\mathrm{N_O}\to\mathrm{N_C}}\left(t, \int_0^t n_{\mathrm{N_O}}(s)\beta_{\mathrm{N}}(V(s)))ds\right) \\ &R_{\mathrm{N_C}\to\mathrm{N_O}}\left(t, \int_0^t n_{\mathrm{N_C}}(s)\alpha_{\mathrm{N}}(V(s)))ds\right), \end{aligned} \tag{3.17}$$

where α_k and β_k depend only on V, and not on the channel state. Thus, at any time t the state $n(t)$ is,

$$\begin{aligned} n(t) = n(0) &+ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} R_{\mathrm{M_O}\to\mathrm{M_C}}(t) + \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} R_{\mathrm{M_C}\to\mathrm{M_O}}(t) \\ &+ \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} R_{\mathrm{N_O}\to\mathrm{N_C}}(t) + \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} R_{\mathrm{N_C}\to\mathrm{N_O}}(t). \end{aligned} \tag{3.18}$$

Here, n represents the number of open or closed channels of each type, R represents an inhomogeneous Poisson process where the first argument is the dependence on time and the second is the mean event rate, and $n(0)$ gives the initial numbers of

"reactants". If we re-express the equations in terms of "concentrations" with units of "reactant" number per area,

$$C(t) = \frac{n(t)}{A}, \tag{3.19}$$

then we may utilize the work of [47] (based on the theoretical work of Kurtz [35, 36, 18]) which shows that in the fixed concentration, large volume limit the continuous time Markov chain description of a chemical reaction network converges to a Gaussian process centered on the mean field ordinary differential equations of the corresponding mass-action kinetics model. In this case, the mean field ordinary differential equations are the deterministic kinetic equations for the Morris-Lecar model. Kurtz's theory of chemical reaction networks along with Fox and Lu's system size expansion approach [24, 23] provide the theoretical foundation which explains why we expect the stochastic Markov chain model at hand to become Gaussian and eventually converge to a deterministic models in an infinite area limit.

So far it has been established that the way currents move per unit area through open channels to determine the voltage is constant as the area increases, but that the Markov chain which controls the number of open channels does change by converging to progressively simpler stochastic descriptions. It is clear now that understanding the evolution of system dynamics as the area increases is fundamentally a mathematical problem. The physics of ion channel movement per unit area across open channel densities is invariant, but probabilistic properties of when a certain density of channels is open to induce those currents is significantly altered when cell membrane area is increased. Now that we have illuminated how scaling alters the structure of the system, a more explicit framework for scaling the model will be outlined.

3.2.3 Scaling Framework

In this section, an attempt will be made to dispel an ambiguity regarding what is happening when the area or number of channels used in the model are increased, and clarify the implications of scaling on other properties of the cell model. The first step is to express the scaling in terms of polar coordinates. Given that the (positive, integer valued) number of channels is equal to the area of the cell times the channel density,

$$N_{\mathrm{M}} = A\rho_{\mathrm{M}}, \tag{3.20}$$

$$N_{\mathrm{N}} = A\rho_{\mathrm{N}}. \tag{3.21}$$

Define,

$$R = \sqrt{\rho_{\mathrm{M}}^2 + \rho_{\mathrm{N}}^2}, \tag{3.22}$$

and

$$\theta = \tan^{-1}\left(\frac{\rho_{\mathrm{M}}}{\rho_{\mathrm{N}}}\right), \tag{3.23}$$

so that any point P in M-N channel number space can be described by,

$$P(A, R, \theta) = AR\,\hat{r} + \theta\,\hat{\theta}, \tag{3.24}$$

for $R > 0$ and $0 \leq \theta \leq \pi/2$.

Since the mean-field current equation is invariant under area scaling, we can think of cells with the same θ and R to be of the same type, i.e. to have the same ratio of the maximal conductances for the two discrete ion channel types, and to have the same total amount of channel noise across the two channel types. Note, that altering the number of channels in a cell model, by changing the channel density, non-uniformly modifies the currents of the model and therefore leads to two distinct cell types. So, a comparison between two members in the same class of cells with different numbers of

channels would require a scaling such that θ and R are constant. Given that density scaling causes non-uniform changes in the model, making it difficult to compare across channel number space, we will replace R with a constant $D = \sqrt{\rho_{\mathrm{M_F}}^2 + \rho_{\mathrm{N_F}}^2}$ so that,

$$P(A, \theta) = AD\,\hat{r} + \theta\,\hat{\theta}, \tag{3.25}$$

and

$$\theta = \tan^{-1}\left(\frac{\rho_{\mathrm{M}}}{\rho_{\mathrm{N}}}\right), \tag{3.26}$$

with the constraint that,

$$\sqrt{\rho_{\mathrm{M}}^2 + \rho_{\mathrm{N}}^2} = D, \tag{3.27}$$

where $\rho_{\mathrm{M_F}} = 60\frac{N}{\mu m^2}$ and $\rho_{\mathrm{N_F}} = 18\frac{N}{\mu m^2}$ (similarly to [28]). So, given a particular number of M and N channels, to find the corresponding area and density in the new coordinate system solve equations 3.20, 3.21, and 3.27 to find,

$$\rho_{\mathrm{M}} = \frac{D}{\sqrt{1 + \left(\frac{N_{\mathrm{N}}}{N_{\mathrm{M}}}\right)^2}}, \tag{3.28}$$

$$\rho_{\mathrm{M}} = D\sqrt{1 - \frac{1}{1 + \left(\frac{N_{\mathrm{N}}}{N_{\mathrm{M}}}\right)^2}}, \tag{3.29}$$

and

$$A = \frac{N_{\mathrm{M}}}{\rho_{\mathrm{M}}}. \tag{3.30}$$

Now the choice of channel densities or channel numbers determines a class of neural models with the same mean-field current equation. The channel configuration plane should then be thought of in terms of similar cell models being scaled in area across lines defined by the ratio of their channel densities.

3.2.4 Stochastic-Deterministic Limits

So far it has been established through Kurtz's theorem that as A becomes large the continuous time Markov chain description of the channel kinetics converges to a Gaussian process centered on the mean field ordinary differential equations of the corresponding mass-action kinetics model. Furthermore, we know that if the channel densities are fixed, $\theta = c$, area scaling only alters the channel state Markov chain. In this section, mixed stochastic-deterministic models are described by taking the infinite area the limit while simultaneously taking θ to 0 or $\pi/2$. Previously, a concentration vector was defined,

$$C(t) = \frac{n(t)}{A} = \begin{bmatrix} \frac{n_{\mathrm{M_O}}}{A} \\ \frac{n_{\mathrm{M_C}}}{A} \\ \frac{n_{\mathrm{N_O}}}{A} \\ \frac{n_{\mathrm{N_C}}}{A} \end{bmatrix}, \tag{3.31}$$

and a large A limit was considered. The limit can be taken in an alternative way by noting that $A = N/\rho$,

$$C(t) = \frac{n(t)}{A} = \begin{bmatrix} \frac{n_{\mathrm{M_O}}\rho_{\mathrm{M}}}{N_{\mathrm{M}}} \\ \frac{n_{\mathrm{M_C}}\rho_{\mathrm{M}}}{N_{\mathrm{M}}} \\ \frac{n_{\mathrm{N_O}}\rho_{\mathrm{N}}}{N_{\mathrm{N}}} \\ \frac{n_{\mathrm{N_C}}\rho_{\mathrm{N}}}{N_{\mathrm{N}}} \end{bmatrix}, \tag{3.32}$$

so that A can be broken into independent parts related to each channel type. Since reactions of each channel type are also isolated from each other, an infinite area limit can be applied to either channel type by taking $N \to \infty$, so that one channel converges to the mean field equations while the other remains stochastic. Note that the effect of taking $N \to \infty$ is realized through equations, 3.28, 3.29, and 3.30. Namely, the area of the cell grows large while the density of one channel increases and the density of the other decreases.

3.2.5 Scaling in the Literature

Previous work analyzing the effect of scaling on neural models differs from this project in a number of ways [57, 10, 59]. Using exact, channel-number tracking [10] or channel-state tracking [59], Markov chain, Hodgkin-Huxley models, prior studies showed how the fractions of open channel dynamics and action potentials changed as the area of the cell increased. By simulating how models of different sizes responded to suprathreshold current stimulation, [10] calculated the mean and variance of the subsequent action potential latency. A similar focus was used by [59], with more attention placed on the convergence of the stochastic model's firing frequency to the classical Hodgkin-Huxley values. In both cases, the densities are fixed and the cell area is scaled. The scaling approach of both [10] and [59] is equivalent to moving across one line in the plane of lines described in section 3.2.3.

In contrast, by varying the area and the densities, we explore the entire plane which represents every channel configuration. An additional distinction between previous simulation studies on scaling and this project is the magnitude of computation. Whereas previous studies have depicted convergence with about 15 points across a single density-line, this project utilizes both efficient error estimation techniques, as well as a powerful numerical implementation of each model on a high performance computing cluster to simulate convergence with a resolution of about 25 points across many lines. Furthermore, scaling simulations are done with multiple variants of the stochastic model.

Chapter 4

The Exact Poisson Region

Simulation of the model defined in section 2.1.2 will be done using an exact channel number tracking algorithm (similar to [10]) which explicitly references the underlying inhomogeneous Poisson processes within the system, and is accurate up to integration error. Pseudocode of the algorithm as well as a visualization of its traces will be given in this chapter, following [1]. As described in section 3.1.1, the exact method becomes slow as the channel numbers increase, and a faster approximation is required. A piecewise constant propensity approximation (PCPA) of the exact algorithm, which is widely used in the literature in place of its exact counterpart, will also be presented along with its pseudocode representation. In [1], Anderson and colleagues showed rapid (weak) convergence of the PCPA to the exact model, in the sense that histograms were indistinguishable under the L1 metric, when channel configurations exceeded 40 channels of each type. Based on these results, we expected that the ISI statistics would also converge under the same conditions. A numerical investigation of the convergence between the two methods will be presented in this chapter and will show that, on the contrary, the ISI statistics do not converge. We discuss the possible implications at the end of this chapter.

4.1 Exact Simulation Algorithm

Since the exact model defined in section 2.1.2 is simply a linear ODE once $n_{\mathrm{M}}(t)$ and $n_{\mathrm{N}}(t)$ are known in equation 2.12, it is clear that the complexity of the simulations arises from generating accurate trajectories of the channel states, or, more specifically, from simulating the trajectories of Poisson processes which have time dependent transition rates. Recall equation 2.7 which gives the number of open channels as a function of time in terms of inhomogenous opening and closing Poisson processes,

$$n(t) = N_{\mathrm{open}}(\, t,\, m_{\mathrm{open}}(t)\,) - N_{\mathrm{close}}(\, t,\, m_{\mathrm{close}}(t)\,), \tag{4.1}$$

where the mean functions are given by,

$$m_{\mathrm{open}}(t) = \int_0^t (\, N - n(s)\,)\, \alpha(\, V(s)\,) ds = \int_0^t \lambda_{\mathrm{open}}(s) ds, \tag{4.2}$$

$$m_{\mathrm{close}}(t) = \int_0^t n(s) \beta(\, V(s)\,) ds = \int_0^t \lambda_{\mathrm{close}}(s) ds. \tag{4.3}$$

A description of the exact simulation algorithm requires a pathwise representation of $n(t)$, which is best understood by first considering a simplified, unit-rate example where $\lambda^*_{\mathrm{open}}(t) = \lambda^*_{\mathrm{close}}(t) = 1$ so that,

$$m^*_{\mathrm{open}}(t) = t, \tag{4.4}$$

$$m^*_{\mathrm{close}}(t) = t. \tag{4.5}$$

Call this simplified channel state process $n^*(t)$, and note that it is described by two homogenous, unit-rate Poisson processes,

$$n^*(t) = N_{\text{open}}(\, t,\, m^*_{\text{open}}(t)\,) - N_{\text{close}}(\, t,\, m^*_{\text{close}}(t)\,) \tag{4.6}$$

$$= N_{\text{open}}(\, t,\, t\,) - N_{\text{close}}(\, t,\, t\,) \tag{4.7}$$

$$= N^*_{\text{open}}(\, t\,) - N^*_{\text{close}}(\, t\,). \tag{4.8}$$

Note that unlike $n(t)$, $n^*(t)$ can be negative, which makes it a nonsensical mathematical model for channel state evolution (but does not stop it from being a simple illustrative example which can be naturally built upon).

The description of the exact algorithm will be carried out in three steps: first a pathwise representation for $n^*(t)$ will be presented, next the representation of $n^*(t)$ will be generalized to describe a pathwise representation of $n(t)$, and lastly the pseudocode used to carry out the exact algorithm on the full two-channel-type model (which takes advantage of the pathwise representation of $n(t)$) will be introduced.

4.1.1 Pathwise Representation of Unit-Rate, Channel-State Processes

Following the presentation found in [1], recall that a unit-rate Poisson process can be constructed from a series of independent exponential random variables with parameter one, $\{e_{\text{open}_\text{i}}\}_{i=1}^{\infty}$. Define,

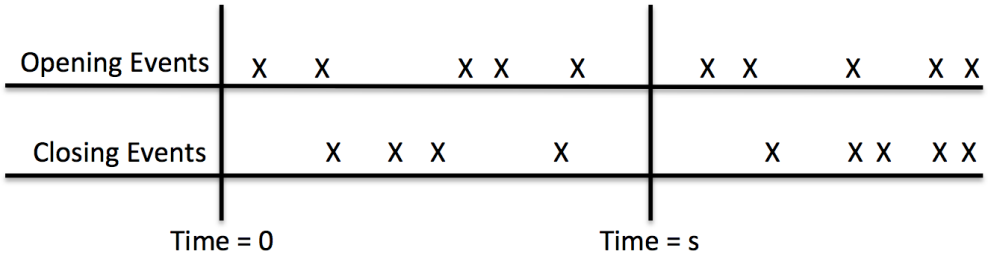
$$\tau_{\text{open}_\text{k}} = \sum_{i=1}^{k} e_{\text{open}_i} \tag{4.9}$$

to be the cumulative sum of these random variables. The value of the unit-rate Poisson process, $N^*_{\text{open}}(t)$, is then the number of points $\{e_{\text{open}_\text{i}}\}_{i=1}^{\infty}$ that come before $t \geq 0$, or the greatest k such that $\tau_{\text{open}_\text{k}} < t$. Similarly, if $\{e_{\text{close}_\text{i}}\}_{i=1}^{\infty}$ is an independent sequence of unit-rate exponential random variables (where each element of $\{e_{\text{close}_\text{i}}\}_{i=1}^{\infty}$

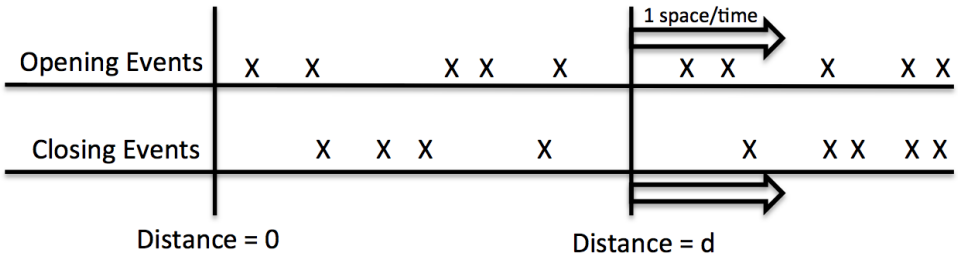
is independent from each element of $\{e_{\text{open}_i}\}_{i=1}^{\infty}$), then,

$$\tau_{\text{close}_k} = \sum_{i=1}^{k} e_{\text{close}_i}, \tag{4.10}$$

and the value of $N^*_{\text{close}}(t)$ is the greatest k such that $\tau_{\text{close}_k} < t$. For example, if X represents each τ_{open_k} and τ_{close_k} in the diagram below, then at time s, $N^*_{\text{open}}(s) = 5$, $N^*_{\text{close}}(s) = 4$, and $n^*(s) = 1$.



In preparation for the next section, there is an alternative interpretation that can be readily generalized to inhomogenous processes: Suppose that the cumulative sums of exponential random variables, τ_{open_k} and τ_{close_k}, are distributed through *space*, and that the second vertical bar moves at a velocity of 1 space unit per time unit (since we have assumed $\lambda^*_{\text{open}}(t) = \lambda^*_{\text{close}}(t) = 1$ for the unit rate processes). The value of each process at time s is then the number of X's that precede $d = \int_0^s 1ds$, or, in the case of $N^*_{\text{close}}(t)$, the greatest k such that $\tau_{\text{close}_k} < \int_0^s 1ds$. This can be expressed more



formally as,

$$n^*(t) = N^*_{\text{open}}\left(\int_0^t 1ds\right) - N^*_{\text{close}}\left(\int_0^t 1ds\right). \tag{4.11}$$

4.1.2 Pathwise Representation of Channel-State Processes

Moving on to the more general case relevant to this project, consider, once again,

$$n(t) = N_{\text{open}}(\, t,\, m_{\text{open}}(t)\,) - N_{\text{close}}(\, t,\, m_{\text{close}}(t)\,), \tag{4.12}$$

where the mean functions are given by,

$$m_{\text{open}}(t) = \int_0^t (\, N - n(s)\,)\, \alpha(\, V(s)\,) ds = \int_0^t \lambda_{\text{open}}(s) ds, \tag{4.13}$$

$$m_{\text{close}}(t) = \int_0^t n(s) \beta(\, V(s)\,) ds = \int_0^t \lambda_{\text{close}}(s) ds. \tag{4.14}$$

and in particular,

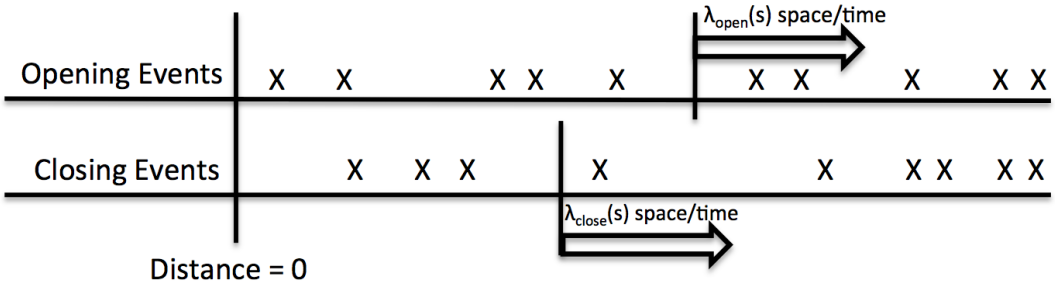
$$\lambda_{\text{open}}(s) = (\, N - n(s)\,)\, \alpha(\, V(s)\,), \tag{4.15}$$

$$\lambda_{\text{close}}(s) = n(s) \beta(\, V(s)\,). \tag{4.16}$$

Extending expression 4.11 to the case of time dependent λ yields a new expression,

$$n(t) = N^*_{\text{open}} \left(\int_0^t \lambda_{\text{open}}(s)\, ds \right) - N^*_{\text{close}} \left(\int_0^t \lambda_{\text{close}}(s)\, ds \right). \tag{4.17}$$

which can be interpreted as follows: There are two independent sequences of unit rate



exponentially distributed random variables, $\{e_{\text{open}_i}\}_{i=1}^{\infty}$ and $\{e_{\text{close}_i}\}_{i=1}^{\infty}$, which can be

distributed across space in the form of two cumulative sums $\tau_{\text{open}_\text{k}} = \sum_{i=1}^{k} e_{\text{open}_\text{i}}$ and $\tau_{\text{close}_\text{k}} = \sum_{i=1}^{k} e_{\text{close}_\text{i}}$. Events occur for each process when their respective "vertical bar", which moves at a velocity of $\lambda_{\text{open}}(s)$ or $\lambda_{\text{close}}(s)$, exceeds $\tau_{\text{open}_\text{k}}$ or $\tau_{\text{close}_\text{k}}$ for some k. More specifically, for all $t \geq 0$, the value of $N^*_{\text{open}}(\int_0^t \lambda_{\text{open}}(s)ds)$ and $N^*_{\text{close}}(\int_0^t \lambda_{\text{open}}(s)ds)$ is the maximal k such that $\tau_{\text{open}_\text{k}} < \int_0^t \lambda_{\text{open}}(s)\, ds$ and $\tau_{\text{close}_\text{k}} < \int_0^t \lambda_{\text{close}}(s)\, ds$, respectively.

4.1.3 Simulating Pathwise Representations of Channel-State Processes

The perspective gained from the last two sections offers a very natural simulation procedure: The X's in the figure above can be readily obtained and the "bar velocities" are known between transitions. All that remains numerically is to integrate the "bar velocities", find the closest τ values which do not exceed the integrals, and then update the counting processes. To generate the X's take two independent sequences of uniform random numbers between 0 and 1, $\{u_{\text{open}_\text{i}}\}_{i=1}^{\infty}$ and $\{u_{\text{close}_\text{i}}\}_{i=1}^{\infty}$, and compute,

$$\tau_{\text{open}_\text{k}} = \sum_{i=1}^{k} -\ln(u_{\text{open}_\text{i}}), \tag{4.18}$$

$$\tau_{\text{close}_\text{k}} = \sum_{i=1}^{k} -\ln(u_{\text{close}_i}). \tag{4.19}$$

The "bar velocities" ($\lambda_{\text{open}}(s)$ and $\lambda_{\text{close}}(s)$) can be found by first noting that between transitions, all parameters are constants except for $V(s)$. Recalling that,

$$\lambda_{\text{open}}(s) = (\, N - n(s)\,)\, \alpha(\, V(s)\,), \tag{4.20}$$

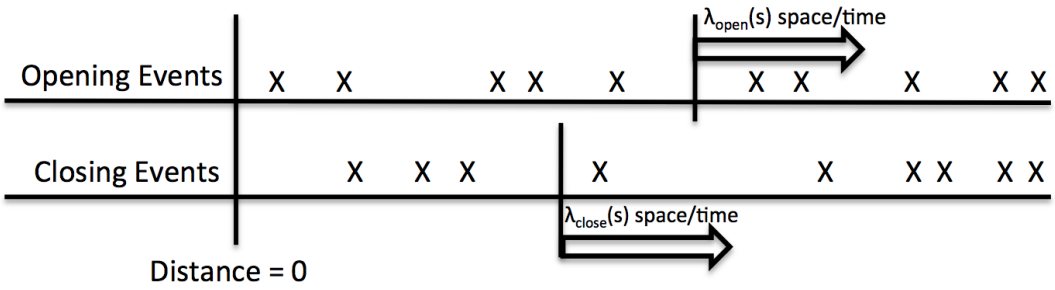
$$\lambda_{\text{close}}(s) = n(s)\beta(\, V(s)\,), \tag{4.21}$$

its clear that between transitions N and $n(s)$ are know constants, and that all parameters of equation 2.12 are known. So, $V(s)$ (and therefore $\lambda_{\text{open}}(s)$ and $\lambda_{\text{close}}(s)$) can be propagated by solving a linear differential equation.

Now that it has been established that τ_{open_k}, τ_{open_k}, $\lambda_{\text{open}}(s)$, and $\lambda_{\text{close}}(s)$ can be obtained between transitions, all that remains is to describe the integration of the "bar velocities", and their comparison to τ_{open_k} and τ_{close_k}. Given an initial state for $n(t)$, as well as τ_{open_k} and τ_{open_k}, the algorithm begins by solving the following equation for the minimal t where,

$$\max\left(\int_0^t \lambda_{\text{open}}(s)\,ds - \tau_{\text{open}_1}, \int_0^t \lambda_{\text{close}}(s)\,ds - \tau_{\text{close}_1}\right) = 0. \tag{4.22}$$

Intuitively, the algorithm uses its knowledge of the velocity of the bars at each point in time to track their position (through integration) and propagates them until one reaches an X. Once a bar reaches an X, or equation 4.22 is solved for t, then $n(t)$ is updated, the relevant τ is incremented and the process repeats. For example, the current state in the figure below is $n(s) = 2$, so the algorithm proceeds by solving the

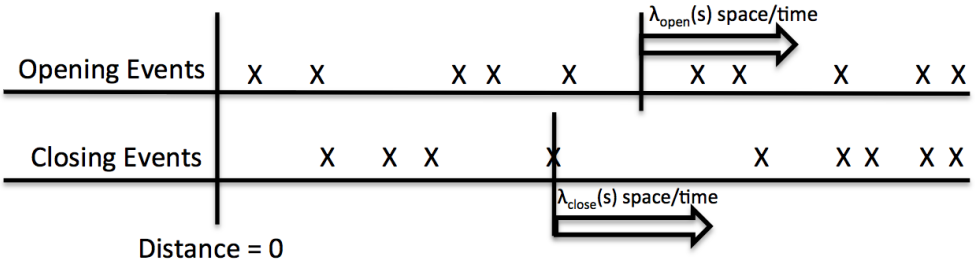


following expression for the minimal t where,

$$\max\left(\int_0^t \lambda_{\text{open}}(s)\,ds - \tau_{\text{open}_6}, \int_0^t \lambda_{\text{close}}(s)\,ds - \tau_{\text{close}_4}\right) = 0. \tag{4.23}$$

Suppose that the system evolved in the way depicted in the figure below to trigger a

closing event. Now $n(s) = 1$ and the algorithm moves forward by solving,



$$\max\left(\int_0^t \lambda_{\text{open}}(s)\,ds - \tau_{\text{open}_6}, \int_0^t \lambda_{\text{close}}(s)\,ds - \tau_{\text{close}_5}\right) = 0. \tag{4.24}$$

What has been presented so far would give an accurate answer, but it is very inefficient since the integrals are calculated from $s = 0$ at each step. If $t_0 \equiv 0 < t_1 < t_2 < ... < t_k < t_{k+1} < ...$ are the transition times of $n(t)$, then we can define variables T_{open_k}, and T_{close_k} to remember integration from the past,

$$T_{\text{open}_1} = 0, \tag{4.25}$$

$$T_{\text{close}_1} = 0, \tag{4.26}$$

and

$$\begin{aligned} T_{\text{open}_k} &= \int_{t_{k-2}}^{t_{k-1}} \lambda_{\text{open}}(s)\,ds + T_{\text{open}_{k-1}}, \\ T_{\text{close}_k} &= \int_{t_{k-2}}^{t_{k-1}} \lambda_{\text{close}}(s)\,ds + T_{\text{close}_{k-1}}. \end{aligned} \tag{4.27}$$

Equation 4.22 can be re-expressed with this new variable as,

$$\begin{aligned} \max\Bigg(&\int_0^{\Delta} \lambda_{\text{open}}(s)\,ds + T_{\text{open}_1} - \tau_{\text{open}_1}, \\ &\int_0^{\Delta} \lambda_{\text{close}}(s)\,ds + T_{\text{close}_1} - \tau_{\text{close}_1}\Bigg) = 0. \end{aligned} \tag{4.28}$$

where the algorithm proceeds by by finding the minimal Δ which solves the equation.

If $q \geq 0$ opening events have occurred, and $r \geq 0$ closing events have occurred the algorithm finds the minimal Δ such that,

$$\max\left(\int_{t_{q+r}}^{\Delta} \lambda_{\text{open}}(s)\,ds + T_{\text{open}_{q+r}} - \tau_{\text{open}_q}, \int_{t_{q+r}}^{\Delta} \lambda_{\text{close}}(s)\,ds + T_{\text{close}_{q+r}} - \tau_{\text{close}_r}\right) = 0. \tag{4.29}$$

The final extension in this section expands the single channel model that has been used thus far (referred to as $n(t)$) to a two channel model which incorporates four independent Poisson processes to describe the opening and closing events of M and N type channels in the Morris-Lecar model defined in section 2.1.2. The extension to two channel types essentially just doubles the number of event types that can happen, and adds two more τ and T parameters. An explicit description of the full algorithm can be seen below.

Exact Markov Chain Algorithm for Stochastic Morris-Lecar Model

The description of the algorithm will begin with an introduction to the abbreviated notation which is not present in the nomenclature section but used in the following outline.

Notation	Description
U	$U[0,1]$ random variable, drawn independently and anew with each mention
τ_{M_O}	Cumulative sum of Exp(1) random variables for M channel opening events, analogous to τ_{open_k}
τ_{N_O}	Cumulative sum of Exp(1) random variables for N channel opening events, analogous to τ_{open_k}
τ_{M_C}	Cumulative sum of Exp(1) random variables for M channel closing events, analogous to τ_{close_k}
τ_{N_C}	Cumulative sum of Exp(1) random variables for N channel closing events, analogous to τ_{close_k}
T_{M_O}	Integrated intensity function for M channel opening events, analogous to T_{open_k}
T_{N_O}	Integrated intensity function for N channel opening events, analogous to T_{open_k}
T_{M_C}	Integrated intensity function for M channel closing events, analogous to T_{close_k}
T_{N_C}	Integrated intensity function for N channel closing events, analogous to T_{close_k}
E_I	Holds the type of the most recent channel transition so that the channel state can be updated later

Algorithm

1. *Initialize*

   $V = V_0,\ M = M_0,\ N = N_0,\ t = 0$

$\tau_{M_O} = -ln(U)$, $\tau_{M_C} = -ln(U)$, $\tau_{N_O} = -ln(U)$, $\tau_{N_C} = -ln(U)$

T_{M_O}, T_{M_C}, T_{N_O}, $T_{N_C} = 0$

2. ***Propagate Exact Solution***

   Input: V, M, N, t, τ_{M_O}, τ_{M_C}, τ_{N_O}, τ_{N_C}, T_{M_O}, T_{M_C}, T_{N_O}, T_{N_C}

   Output: V', t', T'_{M_O}, T'_{M_C}, T'_{N_O}, T'_{N_C}, E_I

3. *Set* $V = V'$, $t = t'$, $T_{M_O} = T'_{M_O}$, $T_{M_C} = T'_{M_C}$, $T_{N_O} = T'_{N_O}$, $T_{N_C} = T'_{N_C}$

4. *Update Channel State Based on* E_I

   if $E_I = E_{M_O}$, *then* $M = M + 1$ *and* $\tau_{M_O} = \tau_{M_O} - ln(U)$

   if $E_I = E_{M_C}$, *then* $M = M - 1$ *and* $\tau_{M_C} = \tau_{M_C} - ln(U)$

   if $E_I = E_{N_O}$, *then* $N = N + 1$ *and* $\tau_{N_O} = \tau_{N_O} - ln(U)$

   if $E_I = E_{N_C}$, *then* $N = N - 1$ *and* $\tau_{N_C} = \tau_{N_C} - ln(U)$

5. *If* $t < t_{\max}$, *then return to 2*

Propagate Exact Solution

1. *Accept Inputs:* V, M, N, t, τ_{M_O}, τ_{M_C}, τ_{N_O}, τ_{N_C}, T_{M_O}, T_{M_C}, T_{N_O}, T_{N_C}

2. *Using* V *as an initial condition as well as constants* M *and* N, *find* $V_s(t)$ *by solving,*

$$C\frac{dV}{dt} = I_{app} - \frac{N}{N_N}G_N(V - V_N) - \frac{M}{N_M}G_M(V - V_M) - G_L(V - V_L)$$

3. *Use non-linear root solver to find the smallest* $0 < \Delta$ *such that,*

$$\max\left(\begin{array}{ll} \int_0^\Delta \alpha_M(V_s(t))\,(N_M - M)dt & + T_{M_O} - \tau_{M_O}, \\ \int_0^\Delta \beta_M(V_s(t))\,M dt & + T_{M_C} - \tau_{M_C}, \\ \int_0^\Delta \alpha_N(V_s(t))\,(N_N - N)dt & + T_{N_O} - \tau_{N_O}, \\ \int_0^\Delta \beta_N(V_s(t))\,N dt & + T_{N_C} - \tau_{N_C} \end{array}\right) = 0$$

4. *Record Which Event Type Caused a Zero Crossing in Event Index,*

   If M Channel opening, then $E_I = E_{M_O}$

If M Channel Closing, then $E_I = E_{M_C}$

If N Channel opening, then $E_I = E_{N_O}$

If N Channel Closing, then $E_I = E_{N_C}$

5. *Update Time and Voltage:* $t' = t + \Delta$, $V' = V_s(\Delta)$

6. *Update Integrated Propensity Variables Using Data From Step 3*

$$
\begin{aligned}
T'_{M_O} &= \int_0^\Delta \alpha_M(V_s(t))\,(N_M - M)dt + T_{M_O} \\
T'_{M_C} &= \int_0^\Delta \beta_M(V_s(t))\,M dt \qquad\quad + T_{M_C} \\
T'_{N_O} &= \int_0^\Delta \alpha_N(V_s(t))\,(N_N - N)dt \;\; + T_{N_O} \\
T'_{N_C} &= \int_0^\Delta \beta_N(V_s(t))\,N dt \qquad\quad + T_{N_C}
\end{aligned}
$$

7. *Produce Outputs:* V', t', T'_{M_O}, T'_{M_C}, T'_{N_O}, T'_{N_C}, E_I

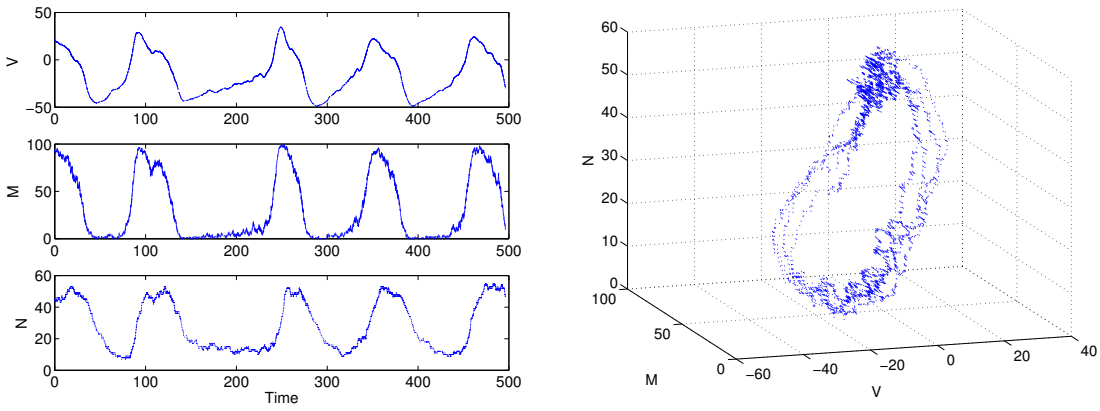


Figure 4.1: Trajectories of the system defined in section 2.1.2, with $N_{\mathrm{N}} = 100$ and $N_{\mathrm{M}} = 100$, using the method outlined in section 4.1. No data is collected in between state transitions, so points are non-uniformly spaced and each point is the result of a channel transition. A comparison with Figure 4.3, which presents the PCPA using the same format, shows that no clear distinction can be made by eye with this channel configuration. Left Panel: Voltage, number of open M channels, and number of open N channels are depicted as a function of time. Right Panel: Same data depicted in state space.

4.2 Piecewise Constant Propensity Approximation (PCPA)

Since the runtime of the previous model is proportional to the total number of channels, an approximation must be developed for large channel number configurations.

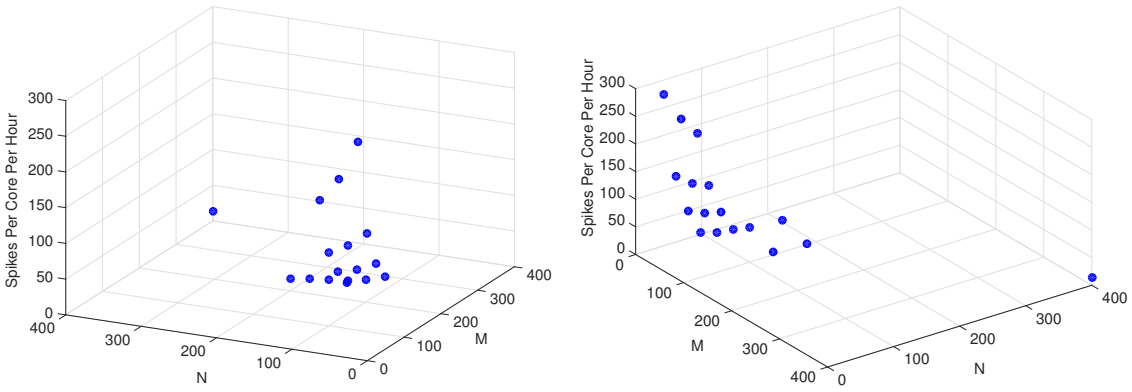


Figure 4.2: The system defined in section 2.1.2 was simulated across 256 cores using the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University with a number of different channel configurations using the algorithm defined in section 4.1.3, and was implemented in MATLAB. The number of cores used for each channel configuration, as well as the resulting number of spikes were both recorded, and the simulation time was approximately 4 hours per core. The point for each channel configuration was calculated by (number of spikes)/(hours per core * cores used). It is clear from the plot that the number of spikes per hour quickly decreases as the total number of channels increases, which motivates the development of faster approximation methods in this section.

The essential observation which facilitates such an approximation is that as the number of channels increases, so do the number of channel events. More channel events per time is equivalent to shorter waiting times between individual events, or that, in some sense, each Δ tends to be smaller. This observation can be understood informally through the total event rate,

$$\begin{aligned} \lambda_{Total}(s) &= \lambda_{\text{open}}(s) + \lambda_{\text{close}}(s) \\ &= (\, N - n(s)\,)\ \alpha(\, V(s)\,) + n(s)\beta(\, V(s)\,), \end{aligned} \tag{4.30}$$

by noting that as N becomes large the voltage trace does not change significantly, $n(s)$ scales with N, and $n(t)$ does not spend more time close to its edges (as seen

in Figure 5.3). These facts are plausible given the discussion in chapter 3, and have been verified numerically. An increase in the event rates as a result of larger channel number configurations acts in equation 4.22 to decrease the resulting solution, Δ,

$$\max\left(\int_0^{\Delta}\lambda_{\text{open}}(s)\,ds-\tau_{\text{open}_1},\int_0^{\Delta}\lambda_{\text{close}}(s)\,ds-\tau_{\text{close}_1}\right)=0. \tag{4.31}$$

Once it is understood that increasing the number of channels tends to reduce the waiting time between events, Δ, it is clear that the form of the integrals present in equation 4.31,

$$\tau_1=\int_0^{\Delta}\lambda(s)ds, \tag{4.32}$$

can be simplified by expanding the upper limit about zero,

$$\tau_1=\lambda(0)\Delta+O\left(\Delta^2\right),\ \text{as}\ \Delta\to 0^+, \tag{4.33}$$

and immediately solved for Δ to first order,

$$\Delta=\frac{\tau_1}{\lambda(0)}+O\left(\Delta^2\right),\ \text{as}\ \Delta\to 0^+. \tag{4.34}$$

If $t_0\equiv 0<t_1<t_2<...<t_k<t_{k+1}<...$ are the transition times of the relevant channel state process, then we can define a point process of transition times, $\Delta_0\equiv 0<\Delta_1<\Delta_2<...<\Delta_k<\Delta_{k+1}<...$, which result from an approximation algorithm, and re-express equation 4.34 as,

$$\Delta_1=\frac{\tau_1}{\lambda(0)}+O\left(\Delta_1^2\right),\ \text{as}\ \Delta_1\to 0^+. \tag{4.35}$$

Moving forward, with the knowledge of Δ_1 the next waiting time would be found by

assuming that Δ_2 tends to be small, and then expanding about the lower limit, Δ_1,

$$\begin{aligned}\tau_2 &= \int_0^{\Delta_2} \lambda(s)ds, \\ &= \int_{\Delta_1}^{\Delta_2} \lambda(s)ds \qquad\qquad + \lambda(0)\Delta_1 + O\left(\Delta_1^2\right), \\ &= \lambda(\Delta_1)\Delta_2 + O\left(\Delta_2^2\right) + \lambda(0)\Delta_1 + O\left(\Delta_1^2\right).\end{aligned} \tag{4.36}$$

Generalizing, with knowledge of all terms of index $i < k$, Δ_k can be found by solving,

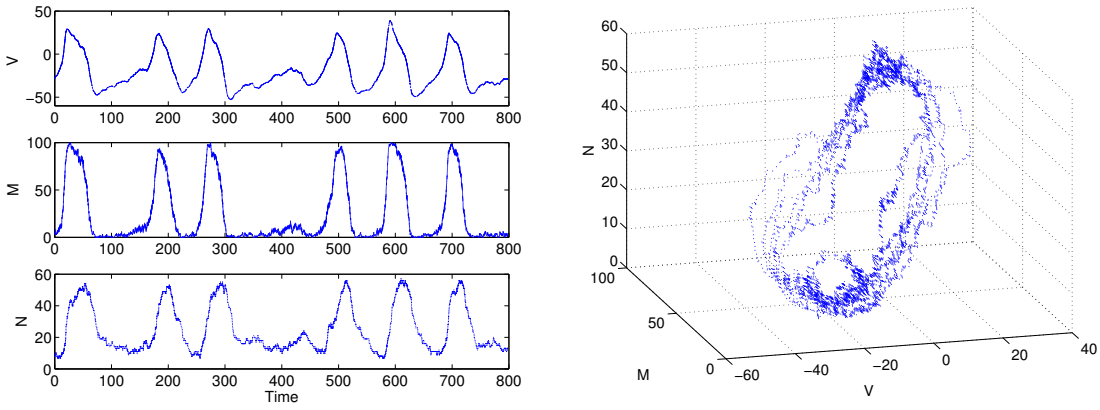


Figure 4.3: Trajectories of the system defined in section 2.1.2, with $N_{\mathrm{N}} = 100$ and $N_{\mathrm{M}} = 100$, using the method outlined in section 4.2. No data is collected in between state transitions, so points are non-uniformly spaced and each point is the result of a channel transition. A comparison with Figure 4.1, which presents the exact model using the same format, shows that no clear distinction can be made by eye with this channel configuration. Left Panel: Voltage, number of open M channels, and number of open N channels are depicted as a function of time. Right Panel: Same data depicted in state space. In this case, a subthreshold voltage oscillation which occurs just after $t = 400$ can be seen as a small circle at the bottom of the larger curve.

$$\begin{aligned}\tau_k &= \int_0^{\Delta_k} \lambda(s)ds, \\ &= \int_{\Delta_{k-1}}^{\Delta_k} \lambda(s)ds + \sum_1^{k-1} \lambda(\Delta_{i-1})\Delta_i + \sum_1^{k-1} O\left(\Delta_i^2\right), \\ &= \sum_1^{k} \lambda(\Delta_{i-1})\Delta_i + \sum_1^{k} O\left(\Delta_i^2\right).\end{aligned} \tag{4.37}$$

By defining,

$$\Delta_{\max} = \max\left(\Delta_1, \ldots, \Delta_k\right), \tag{4.38}$$

the PCPA approximation can be simplified to,

$$\tau_k = \sum_{1}^{k} \lambda(\Delta_{i-1})\Delta_i + O\left(\Delta_{\max}^2\right). \tag{4.39}$$

Terms which have been determined can be separated out of the sum to emphasize the term of interests and to match the form of the integrals in equation 4.29,

$$\tau_k = \lambda(\Delta_{k-1})\Delta_k + \sum_{1}^{k-1} \lambda(\Delta_{i-1})\Delta_i + O\left(\Delta_{\max}^2\right), \tag{4.40}$$

or,

$$\tau_k = \lambda(\Delta_{k-1})\Delta_k + T_{k-1} + O\left(\Delta_{\max}^2\right), \tag{4.41}$$

where T_{k-1} in this context is understood as to be an approximation of equation 4.27 in the last section. Now that equation 4.41 has been derived, it can be noted that the PCPA acts like a forward Euler approximation which would be used to numerically solve ODEs. The pseudo code for the algorithm is included below, note that a notation table is presented in section 4.1.3.

PCPA Markov chain Algorithm

1. *Initialize*

   $V = V_0$, $M = M_0$, $N = N_0$, $t = 0$

   $\tau_{M_O} = -ln(U)$, $\tau_{M_C} = -ln(U)$, $\tau_{N_O} = -ln(U)$, $\tau_{N_C} = -ln(U)$

   T_{M_O}, T_{M_C}, T_{N_O}, $T_{N_C} = 0$

2. ***Propagate PCPA Solution***

   Input: V, M, N, t, τ_{M_O}, τ_{M_C}, τ_{N_O}, τ_{N_C}, T_{M_O}, T_{M_C}, T_{N_O}, T_{N_C}

   Output: V', t', T'_{M_O}, T'_{M_C}, T'_{N_O}, T'_{N_C}, E_I

3. *Set* $V = V'$, $t = t'$, $T_{M_O} = T'_{M_O}$, $T_{M_C} = T'_{M_C}$, $T_{N_O} = T'_{N_O}$, $T_{N_C} = T'_{N_C}$

4. *Update Channel State Based on* E_I

   if $E_I = E_{M_O}$, *then* $M = M + 1$ *and* $\tau_{M_O} = \tau_{M_O} - ln(U)$

   if $E_I = E_{M_C}$, *then* $M = M - 1$ *and* $\tau_{M_C} = \tau_{M_C} - ln(U)$

   if $E_I = E_{N_O}$, *then* $N = N + 1$ *and* $\tau_{N_O} = \tau_{N_O} - ln(U)$

if $E_I = E_{N_C}$, *then* $N = N - 1$ *and* $\tau_{N_C} = \tau_{N_C} - ln(U)$

5. *If* $t < t_{\max}$, *then return to 2*

Propagate PCPA Solution

1. *Accept Inputs:* V, M, N, t, τ_{M_O}, τ_{M_C}, τ_{N_O}, τ_{N_C}, T_{M_O}, T_{M_C}, T_{N_O}, T_{N_C}

2. *Using* V *as an initial condition as well as constants* M *and* N, *find* $V_s(t)$ *by solving,*

$$C\frac{dV}{dt} = I_{app} - \frac{N}{N_N}G_N(V - V_N) - \frac{M}{N_M}G_M(V - V_M) - G_L(V - V_L)$$

3. *Solve for* Δ *such that,*

$$\max\Big(\begin{aligned}&\alpha_M(V)\,(N_M - M)\Delta + T_{M_O} - \tau_{M_O},\\ &\beta_M(V)\,M\Delta + T_{M_C} - \tau_{M_C},\\ &\alpha_N(V)\,(N_N - N)\Delta + T_{N_O} - \tau_{N_O},\\ &\beta_N(V)\,N\Delta + T_{N_C} - \tau_{N_C}\Big) = 0\end{aligned}$$

4. *Record Which Event Type Caused a Zero Crossing in Event Index,*

   If M Channel Opening, then $E_I = E_{M_O}$

   If M Channel Closing, then $E_I = E_{M_C}$

   If N Channel Opening, then $E_I = E_{N_O}$

   If N Channel Closing, then $E_I = E_{N_C}$

5. *Update Time and Voltage:* $t' = t + \Delta$, $V' = V_s(\Delta)$

6. *Update Integrated Propensity Variables*

$$\begin{aligned}T'_{M_O} &= \alpha_M(V)\,(N_M - M)\Delta + T_{M_O}\\ T'_{M_C} &= \beta_M(V)\,M\Delta + T_{M_C}\\ T'_{N_O} &= \alpha_N(V)\,(N_N - N)\Delta + T_{N_O}\\ T'_{N_C} &= \beta_N(V)\,N\Delta + T_{N_C}\end{aligned}$$

7. *Produce Outputs:* V', t', T'_{M_O}, T'_{M_C}, T'_{N_O}, T'_{N_C}, E_I

4.3 Numerical Convergence Between Exact and PCPA Algorithms

To review the discussion in the previous section, recall that as the number of channels is increased the channel event rate, λ, is also increased. A larger event rate causes the waiting times between events, Δ, to be smaller, which causes the second order term on the right hand side of,

$$\int_0^{\Delta_k} \lambda(s)ds = \sum_1^k \lambda(\Delta_{i-1})\Delta_i + O\left(\Delta_{\max}^2\right), \tag{4.42}$$

to become less prominent. Since the left hand side is the fundamental equation of the exact model and the right hand side is from the PCPA model, it is then natural to expect the PCPA to converge to the exact model for sufficiently large channel numbers. In the following discussion, the convergence of these two models in terms of the ISI statistics will be evaluated numerically.

In addition to the informal argument given above, [1] provides indirect evidence of the convergence between the exact and PCPA algorithms. They do so by simulating the same system this project defines in section 2 through a number of cycles using the exact and PCPA algorithms, and then collecting the states of each into two numerical histograms. The L_1 difference between the two histograms was shown to become small as the number of channels is increased.

More specifically, simulation using each method gives an array containing S samples of voltage, M-channel, and N-channel states that the system occupies throughout its evolution. If each point (V, M, N) from these arrays is thought of as a point in state space (as depicted in Figure 4.3), one can ask, given the simulation data, what is the probability that the system will enter into a particular cube of state space? In [1], 100 voltage bins are made through an equal partition of the values between the max and min voltage values, so that the cubes are defined by each discrete channel

state and the voltage bins. Once the cubes are defined, the numerical histogram, H, is found by counting the number of the points in each cube, and the L_1 difference between two histograms is defined as

$$L_1(H_1 - H_2) = \frac{\sum_{(V_b,M,N)} |H_1(V_b, M, N) - H_2(V_b, M, N)|}{S} \tag{4.43}$$

where V_b denotes a voltage bin, M and N denote the number of open channels, $H_k(V_b, M, N)$ is the number of points in the cube defined by (V_b, M, N) from histogram k, and each numerical histogram is understood to contain a total of S points. While the work of [1] shows that the L_1 difference between numerical histograms of the exact and PCPA algorithms is small after about 40 channels of each type are used, it does not immediately follow that the ISI mean, variance, or CV which result from the two algorithms also converge. To see this, note that two stochastic processes can have the same stationary distribution (the histogram) while having different pathwise properties, such as mean return times between surfaces. While [1] showed numerical evidence of weak convergence (convergence of the histograms of the stationary distributions) the authors also demonstrated lack of strong convergence, in that trajectories generated by the exact and PCPA algorithms respectively, and driven with identical Poisson process inputs, diverged after only a few spikes.

An evaluation of the numerical convergence of the ISI distribution between the two algorithms contained in Figure 4.4 suggests that the two algorithms do not visibly converge in a region suggested by the results of [1]. In particular, note that the mean ISI is systematically higher in the approximation. While this result is unexpected, what is certain is that the convergence of the PCPA model to the SDE model introduced in the next section was in accordance with expectation. So, there are three possibilities: a) the exact model was not implemented correctly in this project b) the results of [1] are incorrect or misleading c.) histogram convergence seen in [1] does

not imply convergence of ISI mean, variance, or CV. One way to test (a) is to attempt to replicate the histogram convergence results seen in [1] with this project's model implementations. If the convergence results were reproduced the natural conclusion is (c). Further research on the relationship between histogram convergence and ISI distribution convergence could also be successful in illuminating this mystery.

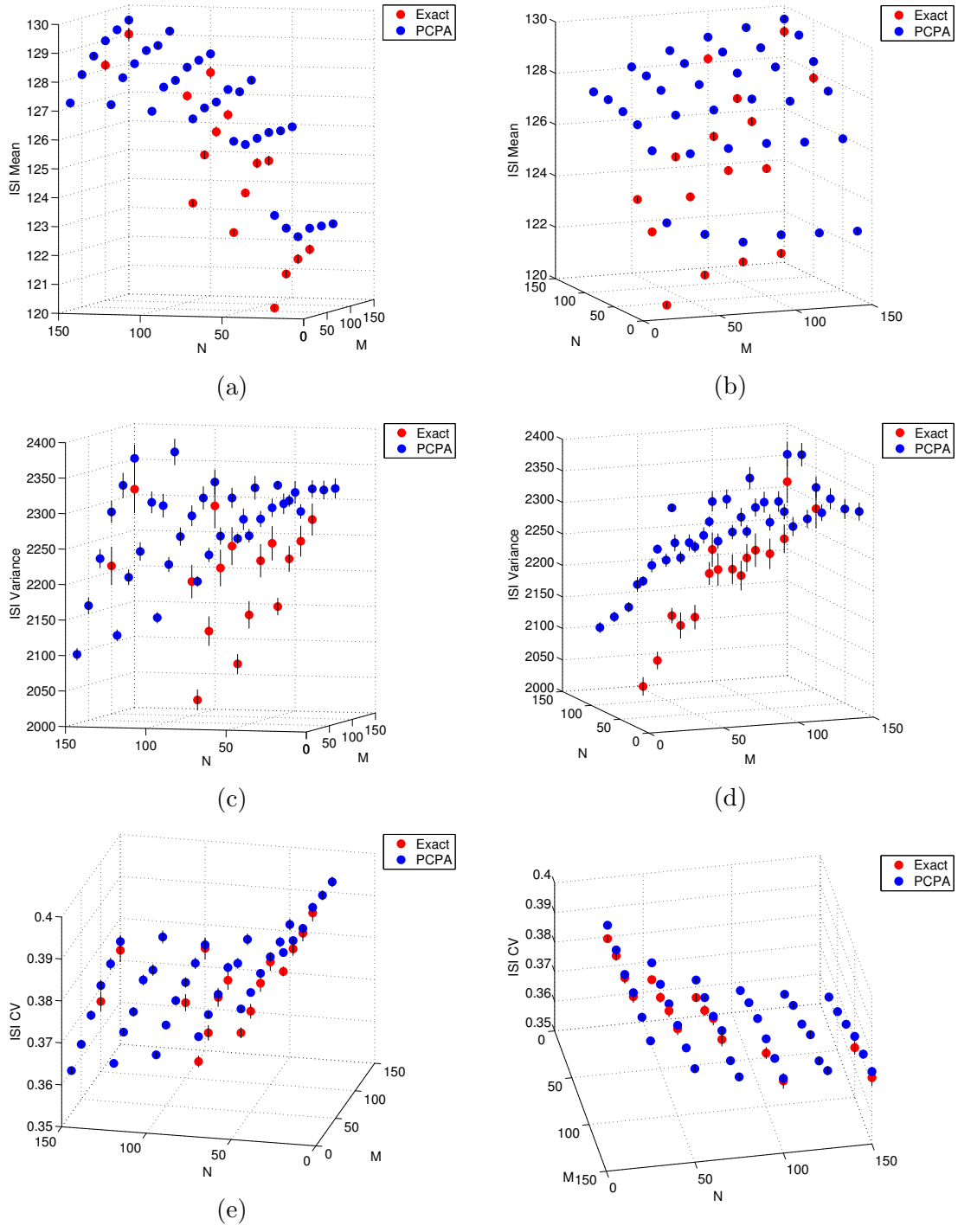


Figure 4.4: The figure shows the ISI mean, ISI variance and ISI CV for the system defined in section 2.1.2, simulated using the exact (red, section 4.1) and PCPA (blue, section 4.2) algorithms with various channel configurations. The ISI distribution of each of the channel configurations from each simulation method is found from the associated voltage trace (using a dual-threshold crossing method as described in section 2.1.3) and enables calculation of the ISI mean (a,b), variance (c,d), and CV (e,f). The standard error of each statistical measure is represented by black lines (often too small to be visible) and is obtained using approximations outlined in section 7.2. Note that each column offers a different view of the same figure, and that there are more PCPA (blue) points because the faster method makes them significantly easier to simulate. Contrary to the convergence argument given at the beginning of the section, and the results of [1], which show histogram convergence after channels numbers exceed 40 of each type, ISI statistics generated by the exact algorithm and the PCPA algorithm are distinct across the entire simulated surface. The absence of convergence between the ISI statistics is quantified in Figure 4.5.

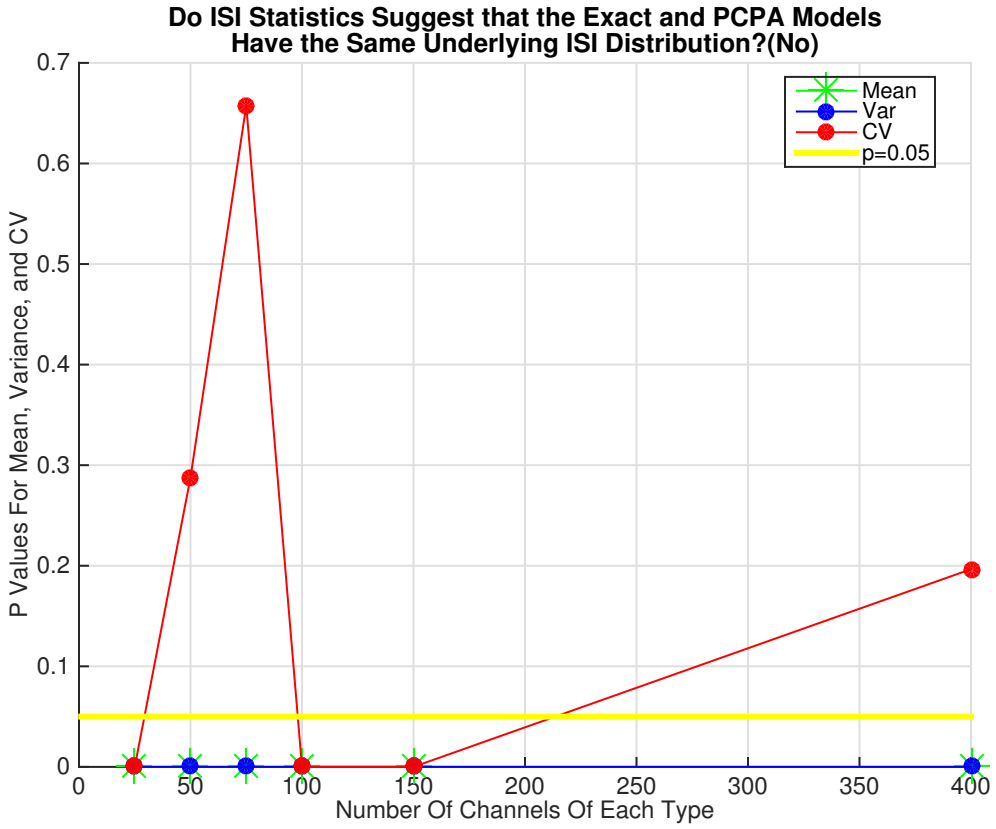


Figure 4.5: If N is the number of channels of each type, then under the null hypothesis that $H_0 : \mathrm{ISI}_N^{\mathrm{Exact}} = \mathrm{ISI}_N^{\mathrm{PCPA}}$, this plot shows the probability of observing values equal to or more extreme than the measured difference between the ISI mean, variance, and CV which result from the exact and PCPA models (as depicted in Figure 4.4 with the addition of one point at $N = 400$). The yellow line gives the standard significance level of 0.05, and clarifies that the measured ISI statistics do not provide enough evidence to support the null hypothesis H_0. We can conclude that the convergence of the exact model to the PCPA model is not statistically significant along the $N_N = N_M$ line (See section 4.3 for discussion). The hypothesis test was developed in section 7.4 and is given in equation 7.57 which uses the test statistic from equation 7.58. Note that it is possible for the mean and variance which result from each of the models to be far apart, but for the CVs to be close (since the CV is a standardized measure). See figure 5.5 for an example of a successful test.

Chapter 5

The Gaussian Region

Though the PCPA provides a significant speed increase over the exact model, the fundamental structure remains the same, and the runtime of the algorithm scales with the number of channels used in the cell configuration. Increasing the channel numbers increases the runtime of the exact and PCPA algorithms because the simulations must stop after each channel transition to propagate the solution. As the number of channels increases, the waiting time between events grows small, and the algorithms are forced to stop and propagate the solution over very small time steps. The PCPA defined in the previous section made each step less costly by substituting a constant value in place of the integral used in the exact algorithm. However, the time-step bottleneck persists, and makes the PCPA approximation prohibitively slow when channel numbers exceed a few thousand.

In this section an approximation of the model defined in section 2.1.2 is formulated in a stochastic differential equation (SDE), valid when event rates, λ, are large enough for the Poisson processes to be approximated by Gaussian processes. The key to the SDE model is that for a given time step the SDE model approximates all of the effects of noise which have occurred over that time step into a single Gaussian random variable, thereby eliminating any of the time step bottlenecks present in the exact

and PCPA models. As a result, the runtime of the SDE model is independent of the channel configuration.

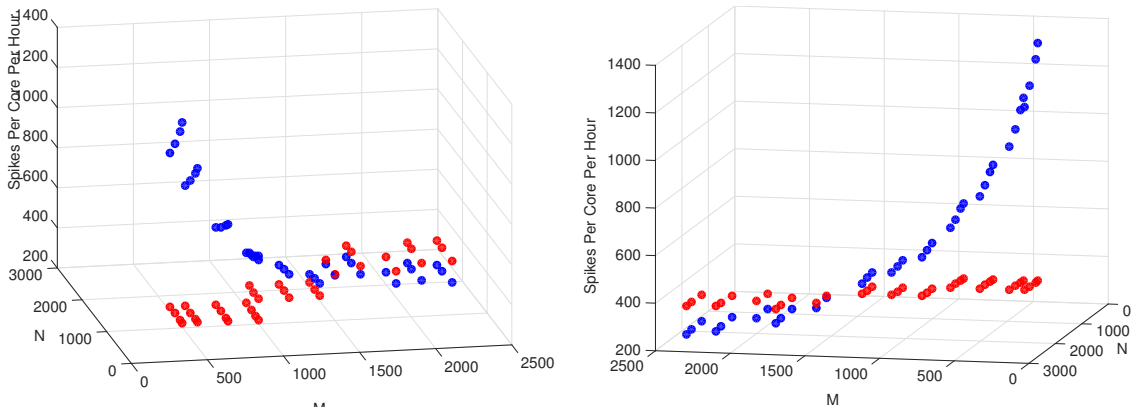


Figure 5.1: The system defined in section 2.1.2 was simulated across 256 cores using the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University with a number of different channel configurations using the algorithm defined in section (blue, 4.2) and (red, 5.2), both implemented in MATLAB. The number of cores used for each channel configuration, as well as the resulting spikes were recorded, and the simulation time was approximately 3 hours per core. The point for each channel configuration was calculated by (number of spikes)/(hours per core * cores used). It is clear from the plot that for the PCPA (blue) points the number of spikes per hour quickly decreases as the total number of channels increases, which motivates the development of faster approximation methods in this section, indicated by the red points. The fundamental bottleneck in the algorithms defined in section 4.1.3 and 4.2 is that they stop at each channel transition, which forces them to operate on a progressively finer grid as the density of channel transitions increases. As shown with the red dots, the methods defined in this chapter yield an SDE which approximates a number of channel transitions using a vector of Gaussian random variables at each fixed time step, and has a constant number of spikes per core per hour (which depends on the timestep, $\Delta t = 0.0043$). After about 1500 channels of each type are used, the SDE model becomes faster, and produces spikes at a constant rate as the channel numbers are further increased.

Once the new SDE model is derived, its convergence to the PCPA model in terms of ISI statistics will be evaluated numerically. After establishing that the SDE model is valid though its convergence to the PCPA model, its speed will be utilized to simulate the ISI mean, variance, and CV surfaces across large sections of the channel configuration space.

5.1 Gaussian Approximation of PCPA Algorithm

Beginning with the Exact Hybrid Markov model for the Morris-Lecar Neuron model defined in section 2.1.2,

$$\begin{aligned} C\frac{dV}{dt} = I_{\text{app}} &- \frac{n_{\text{M}}(t)}{N_{\text{M}}}G_{\text{M}}(V - V_{\text{M}}) \\ &- \frac{n_{\text{N}}(t)}{N_{\text{N}}}G_{\text{N}}(V - V_{\text{N}}) - G_{\text{L}}(V - V_{\text{L}}), \end{aligned} \tag{5.1}$$

we will show that when N_{M} and N_{N} are large this equation can be approximated by a SDE model. To start, define what will soon become continuous stochastic variables,

$$X(t) = \frac{n_{\text{M}}(t)}{N_{\text{M}}}, \qquad Y(t) = \frac{n_{\text{N}}(t)}{N_{\text{N}}}. \tag{5.2}$$

So that now,

$$\begin{aligned} C\frac{dV}{dt} = I_{\text{app}} &- X(t)G_{\text{M}}(V - V_{\text{M}}) \\ &- Y(t)G_{\text{N}}(V - V_{\text{N}}) - G_{\text{L}}(V - V_{\text{L}}). \end{aligned} \tag{5.3}$$

The focus of the derivation will now be placed on $X(t)$ for simplicity, since the procedure for $Y(t)$ is largely the same. Note that by equation 2.7,

$$X(t) = \frac{N_{\text{open}}(\, t,\, m_{\text{open}}(t)\,) - N_{\text{close}}(\, t,\, m_{\text{close}}(t)\,)}{N_{\text{M}}}. \tag{5.4}$$

Moving forward, the derivation will focus further on the first term above, since the procedure for the second term is similar. Begin by naming the first term for convenience,

$$X_{\text{open}}(t) = \frac{N_{\text{open}}(\, t,\, m_{\text{open}}(t)\,)}{N_{\text{M}}}, \tag{5.5}$$

where the dependence on N_{M} will be suppressed. The approach will be to show that for sufficiently large channel numbers, N_{M}, the mean event rate, m_{open}, which governs the number of channels that open in a small time interval, is large enough for the

Poisson process to be adequately approximated by a Gaussian process. Once the increments are expressed in terms of a Gaussian process, a limit is taken and the associated SDE model comes out naturally.

Start with an increment of X_{open} which gives the number of channel opening events at time t over a small interval of length Δt,

$$\begin{aligned} &X_{\text{open}}(t+\Delta t) - X_{\text{open}}(t) \\ &\quad = \frac{1}{N_{\text{M}}} \left[N_{\text{open}}(t+\Delta t, m_{\text{open}}(t+\Delta t)) - N_{\text{open}}(t, m_{\text{open}}(t)) \right]. \end{aligned} \tag{5.6}$$

For a fixed t, the expression above is a Poisson random variable distributed with an event parameter,

$$\begin{aligned} \lambda_{t,t+\Delta t} &= \int_t^{t+\Delta t} (\, N_{\text{M}} - n(s)\,)\, \alpha_{\text{M}}(\, V(s)\,) ds \\ &= N_{\text{M}} \int_t^{t+\Delta t} (\, 1 - X(s)\,)\, \alpha_{\text{M}}(\, V(s)\,) ds. \end{aligned} \tag{5.7}$$

For sufficiently small Δt we can make a piecewise constant propensity approximation, as was done in section 4.2, to simplify the expression,

$$\lambda_{t,t+\Delta t} = N_{\text{M}}(\, 1 - X(t)\,)\, \alpha_{\text{M}}(\, V(t)\,) \Delta t + O\left(\Delta t^2\right). \tag{5.8}$$

Stepping back for a moment, we have shown that for small Δt,

$$X_{\text{open}}(t+\Delta t) - X_{\text{open}}(t) \sim \text{Poisson}\left(\lambda_{t,t+\Delta t}\right), \tag{5.9}$$

and would like to approximate this random variable with a normal random variable distributed as $\mathcal{N}\left(\lambda_{t,t+\Delta t}, \lambda_{t,t+\Delta t}\right)$. A standard rule of thumb is that the approximation is "good" for $\lambda > 20$, so let's find some conservative values for each parameter in 5.8 to see if the approximation is reasonable. Figure 5.3 shows that for a system with 1000 channels of each type, most of the time the number of open channels is less than

900. So, a conservative value of $(1 - X(t))$ is 0.1. The rate function, depicted in Figure 2.1, is large when $(1 - X(t))$ is small, but 0.1 seems like a good conservative value. Lastly, let's use 0.01 for the time step Δt. With these values we have,

$$\begin{aligned} \lambda_{t,t+\Delta t} &= N_{\mathrm{M}}(\,1 - X(t)\,)\,\alpha_{\mathrm{M}}(\,V(t)\,)\Delta t + O\left(\Delta t^2\right) \\ 20 &= N_{\mathrm{M}}(0.1)(0.1)(0.01) + O\left(0.01^2\right). \end{aligned} \tag{5.10}$$

Therefore, a conservative estimate suggests that the approximation should be valid when the number of M channels exceeds $N_{\mathrm{M}} \approx 20000$. The calculation for the other terms in $X(t)$ and $Y(t)$ is similar, so we will proceed with the assumption that Poisson $(\lambda_{t,t+\Delta t})$ is well approximated by $\mathcal{N}(\lambda_{t,t+\Delta t}, \lambda_{t,t+\Delta t})$. As discussed in section 5.3, this project does not directly calculate the error of the Gaussian approximation, but instead numerically searches for channel configurations where the Gaussian approximation and the PCPA agree in terms of ISI statistics (section 5.3), from which it can be concluded that the Gaussian approximation is sound (or not).

Following this train of thought,

$$\begin{aligned} &X_{\mathrm{open}}(t + \Delta t) - X_{\mathrm{open}}(t) \\ &\quad\approx \frac{1}{N_{\mathrm{M}}}\left[\mathcal{N}\left(\lambda_{t,t+\Delta t}, \lambda_{t,t+\Delta t}\right)\right] \\ &\quad= \mathcal{N}\left(\frac{\lambda_{t,t+\Delta t}}{N_{\mathrm{N}}}, \frac{\lambda_{t,t+\Delta t}}{N_{\mathrm{N}}^2}\right) \\ &\quad= \mathcal{N}\left((1 - X(t))\,\alpha_{\mathrm{M}}(\,V(t)\,)\Delta t, \frac{(1 - X(t))\,\alpha_{\mathrm{M}}(\,V(t)\,)\Delta t}{N_{\mathrm{N}}}\right), \end{aligned} \tag{5.11}$$

where terms in $O\left(\Delta t\right)^2$ have been neglected to simplify presentation, and N_{M} is understood to be sufficiently large (determined in section 5.3). Taking the limit as $\Delta t \to 0$ gives an SDE which governs the behavior of X_{open},

$$dX_{\mathrm{open}} = (1 - X(t))\,\alpha_{\mathrm{M}}(\,V(t)\,)dt + \sqrt{\frac{(1 - X(t))\,\alpha_{\mathrm{M}}(\,V(t)\,)}{N_{\mathrm{M}}}}dW_{X_1}(t). \tag{5.12}$$

The SDE for X_{close} takes a similar form,

$$dX_{\text{close}} = X(t)\,\beta_{\text{M}}(\,V(t)\,)dt + \sqrt{\frac{X(t)\,\beta_{\text{M}}(\,V(t)\,)}{N_{\text{M}}}}dW_{X_2}(t), \tag{5.13}$$

where dW_{X_1} is independent of dW_{X_2}. Define,

$$\mu_X\,(V,X) = (1-X)\,\alpha_{\text{M}}(\,V\,) - X\,\beta_{\text{M}}(\,V\,), \tag{5.14}$$

$$\sigma_X\,(V,X) = \sqrt{(1-X)\,\alpha_{\text{M}}(\,V\,) + X\,\beta_{\text{M}}(\,V\,)}, \tag{5.15}$$

$$\epsilon_X = \frac{1}{\sqrt{N_{\text{M}}}}, \tag{5.16}$$

and note that the difference of independent normal random variables is normal. These observations yield the full SDE,

$$dX_{\text{open}} - dX_{\text{close}} = dX = \mu_X(V,X)dt + \epsilon_X\sigma_X(V,X)dW_X. \tag{5.17}$$

Given dW_Y, which is independent of dW_X, the SDE for $Y(t)$ takes exactly the same form,

$$dY = \mu_Y(V,X)dt + \epsilon_Y\sigma_Y(V,X)dW_Y, \tag{5.18}$$

except,

$$\mu_Y\,(V,Y) = (1-Y)\,\alpha_{\text{N}}(\,V\,) - Y\,\beta_{\text{N}}(\,V\,), \tag{5.19}$$

$$\sigma_Y\,(V,Y) = \sqrt{(1-Y)\,\alpha_{\text{N}}(\,V\,) + Y\,\beta_{\text{N}}(\,V\,)}, \tag{5.20}$$

and,

$$\epsilon_Y = \frac{1}{\sqrt{N_{\text{N}}}}. \tag{5.21}$$

Hence, the full system can be written as,

$$\begin{bmatrix} dV \\ dX \\ dY \end{bmatrix} = \begin{bmatrix} \mu_V(V,X,Y) \\ \mu_X(V,X,Y) \\ \mu_Y(V,X,Y) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \epsilon_X \sigma_X(V,X,Y) dW_X(t) \\ \epsilon_Y \sigma_Y(V,X,Y) dW_Y(t) \end{bmatrix}. \tag{5.22}$$

where from equation 5.3 it is clear that,

$$\begin{aligned} \mu_V(V,X,Y) = \frac{1}{C}\Big[& I_{\text{app}} - X(t) G_{\text{M}}(V - V_{\text{M}}) \\ & - Y(t) G_{\text{N}}(V - V_{\text{N}}) - g_{\text{L}}(V - V_{\text{L}}) \Big]. \end{aligned} \tag{5.23}$$

Note that since all the noise in the model arises from the random gating of ion channels, there is no noise term directly perturbing V.

5.2 SDE Simulation Algorithm

Now that the system has been re-expressed in the form of an SDE in equation 5.22, simulation is just a straightforward application of the Euler-Maruyama method for solving systems of SDEs. More specifically, the stochastic channel equations defined by equation 5.17 and 5.18 are propagated forward using a fixed Euler-Maruyama step. Since μ_V defines a deterministic, linear, ODE once X and Y are defined, it can be solved, and the solution can be propagated exactly over each time step. As a result, integration error comes exclusively from the Euler-Maruyama method which has $O\left(\sqrt{\Delta t}\right)$ error. For small channel numbers, additional error is introduced when the Euler-Maruyama solution for the channel states X and Y are occasionally calculated to be greater than 1 or less than zero (which are senseless values). To correct for a upper or lower crossing, the channel states are artificially set back to a well-defined values 1 or 0, respectively. The probability of a boundary error diminishes

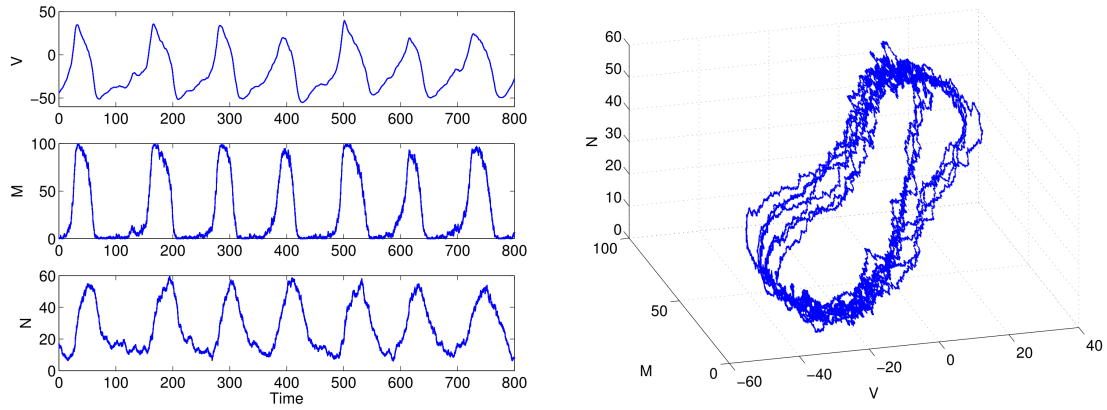


Figure 5.2: Trajectories of the system defined by equation 5.3, with $N_{\mathrm{N}} = 100$ and $N_{\mathrm{M}} = 100$, are simulated using an Euler-Maruyama based method outlined in section 5.2 with a fixed time step of $\Delta t = 0.0043$ (which is proportional to the average time step of the PCPA model for channel configurations around 2000). The channel values N and M are continuous unlike the other models, and are derived from $M = N_{\mathrm{M}}X(t)$ and $N = N_{\mathrm{N}}Y(t)$ for comparison purposes. The model defined by equation 5.3 was derived under the assumption that the channel states were far from their maximum and minimum values. The evolution of the channel states depicted in the lower two subfigures show that this assumption holds for the N channel trace but that the M channel trace frequently rests in a minimum region around 0. Furthermore, this trace required 1333 lower bound corrections of the X variable (step 3 in pseudocode given in section 5.2) and 11 upper bound corrections. While the SDE method was not established to be accurate for this channel configuration, this figure facilitates a comparison to Figures 4.3 and 4.1. It is immediately clear that the time step for the Euler-Mayruyama method is much smaller than the effective time step for the methods in chapter 4. However, the dynamics appear to be about the same. A comparison with Figure 5.3 shows that as the channel numbers are increased the traces are smoothed and the M channel trace is lifted out of the minimal region in accordance with the assumption of the approximation. Left Panel: voltage, number of open M channels ($M = N_{\mathrm{M}}X(t)$), and number of open N channels ($N = N_{\mathrm{N}}Y(t)$) are depicted as a function of time. Right Panel: Same data depicted in state space.

as the number of channels increases because the system spends less time close to the edges, as is shown through a comparison of Figures 5.2 and 5.3.

Stochastic Differential Equation Algorithm

1. *Initialize*

$$k = 1,\ t_k = 0,\ V_k = V_0,\ X_k = X_0,\ Y_k = Y_0,$$

2. *Propagate X and Y equations with Euler Mayruyama method*

$$\begin{aligned} X_{k+1} &= X_k + \big(\alpha_M(V_k)(1 - X_k) - \beta_M(V_k)X_k\big)\Delta t \\ &\quad + \left(\sqrt{\frac{\alpha_M(V_k)(1 - X_k) - \beta_M(V_k)X_k}{N_M}}\right)\Delta W_k \\ Y_{k+1} &= Y_k + \big(\alpha_N(V_k)(1 - Y_k) - \beta_N(V_k)Y_k\big)\Delta t \\ &\quad + \left(\sqrt{\frac{\alpha_N(V_k)(1 - Y_k) - \beta_N(V_k)Y_k}{N_N}}\right)\Delta W'_k \end{aligned}$$

where $\Delta W_k = Z\sqrt{\Delta t}$, $\Delta W'_k = Z'\sqrt{\Delta t}$, and $Z, Z' \sim N(0,1)$ so that $Z \perp\!\!\!\perp Z'$.

3. *Artificiality correct any erroneous barrier crossings*

$$\begin{aligned} &\textit{If } X_{k+1} > 1, X_{k+1} = 1 \\ &\textit{If } X_{k+1} < 0, X_{k+1} = 0 \\ &\textit{If } \ Y_{k+1} > 1, \ Y_{k+1} = 1 \\ &\textit{If } \ Y_{k+1} < 0, \ Y_{k+1} = 0 \end{aligned}$$

4. *Using V_k as an initial condition as well as constants X_k and Y_k, solve*

$$\begin{aligned} C\frac{dV}{dt} = I_{app} &- Y_k G_N(V - V_N) \\ &- X_k G_M(V - V_M) - G_L(V - V_L) \end{aligned}$$

   to find $V_{k+1} = V(t_k + \Delta t)$

5. *Set $k = k + 1$*

6. *If $t_k < t_{\max}$, then return to 2*

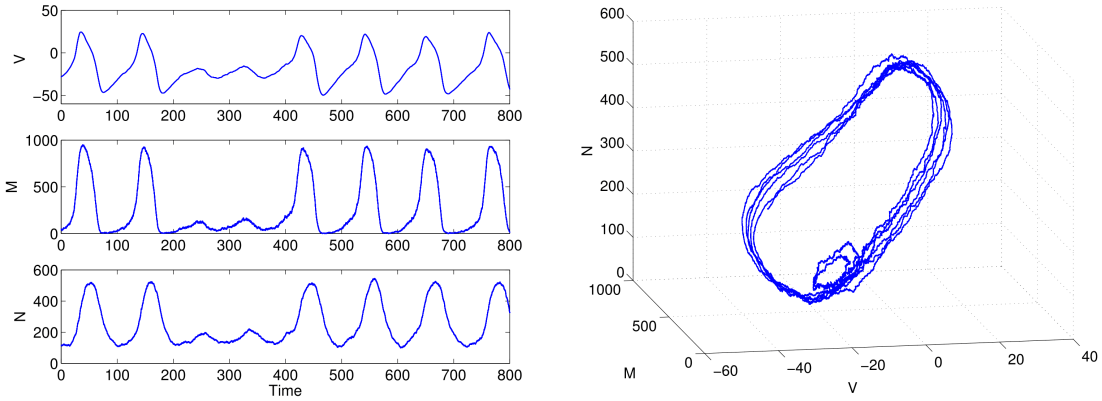


Figure 5.3: Trajectories of the system defined by equation 5.3, with $N_{\mathrm{N}} = 1000$ and $N_{\mathrm{M}} = 1000$, using a Euler-Maruyama based method outlined in section 5.2 with a fixed time step of $\Delta t = 0.0043$ (which is proportional to the average time step of the PCPA model for channel configurations around 2000). The channel values N and M are continuous unlike the other models, and are derived from $M = N_{\mathrm{M}} X(t)$ and $N = N_{\mathrm{N}} Y(t)$ for comparison purposes. The model defined by equation 5.3 was derived under the assumption that the channel states were far from their maximum and minimum values. The evolution of the channel states depicted in the lower two subfigures shows that this assumption holds for the N channel trace, and that (unlike the 100 channel example in Figure 5.2) the M trace only briefly dips into a minimal region around 0. While the 100 channel configuration trace required 1333 lower and 11 upper bound corrections to the X variable (step 3 in pseudocode given in section 5.2), this configuration only required 8 lower and 0 upper bound corrections to the X variable. A comparison to Figure 5.2 shows the increased number of channels in this model cause the trace to smooth, and places it on a more regular path. However, the lower right side of the of right panel and the traces between 200 and 400 time units on the left plane illustrate that occasionally the is system perturbed enough to move far from the deterministic path to great effect. Left Panel: voltage, number of open M channels ($M = N_{\mathrm{M}} X(t)$), and number of open N channels ($N = N_{\mathrm{N}} Y(t)$) are depicted as a function of time. Right Panel: same data depicted in state space.

5.3 Numerical Convergence of PCPA and Stochastic Differential Equation Model

Beginning with the model defined in section 4.1, which simulates the system defined in section 2.1.2 up to integration error, this project set out to establish the convergence of the exact model to its PCPA. Though inconsistencies in the exact model were discovered (and discussed in section 4.3), the project continues under the assumption that the PCPA model should converge to the exact model when channel numbers exceed about 50, in accordance with the indirect results of [1]. Under this assumption, the PCPA is considered to be numerically equivalent to the exact model for channel numbers in the acceptable region. By extension, if numerical equivalence can be established between the PCPA and SDE model, then the SDE model would be validated, and could be utilized as a fast simulation technique for large channel configurations (since its runtime is independent of channel numbers). The following presentation establishes the numerical equivalence between these two algorithms and utilizes the speed of the newly validated SDE model to depict the ISI mean, variance, and CV surfaces.

There are four potential sources of discrepancy between the PCPA and SDE model. Each will presented and then acknowledged after a ":".

(1) The correction of boundary error which occurs when the Euler-Maruyama method propagates the X or Y variables above 1 or below zero: as illustrated in Figure 5.2 which had 1333 lower and 11 upper bound corrections, and Figure 5.3 which only had 8 boundary error crossings, the error becomes negligible as the channel numbers increase.

(2) The error terms in $O\left(\Delta t\right)^2$, which originate from the PCPA in equation 5.8, persist implicitly up to the final equation, 5.22: these terms can be made negligible with a sufficiently small time step, or made consistent with the PCPA model by

making the SDE time step, $\Delta t = 0.0043$, close to the average time step of the PCPA model ($\mathbf{E}[\Delta t] \approx 0.005$ for channel configurations around 2000).

(3) There is a difference in the miscellaneous numerical error which results from fundamentally different simulation methods: this project assumes that this difference is negligible.

(4) There is error in the Gaussian approximation of Poisson process: error in the approximation of the Poisson distribution by the Gaussian as a function of channel number is not directly quantified in this project. Instead of directly calculating the error in the Gaussian approximation, this project finds the sufficient configurations indirectly by simulating a variety with the SDE model and determining which converge to their associated configuration in the (indirectly [1]) validated PCPA model.

Results of Figure 5.4 show that for all channel configurations which exceed 400 channels of each type, the PCPA and SDE models are within a standard deviation of each other in terms of the mean, variance, and CV. Furthermore, Figure 5.5 quantitatively validates the convergence of the SDE model to the PCPA model along the $N_N = N_M$ line. Though there is some uncertainty regarding what happens farther away from the $N_N = N_M$ line in Figure 5.4, it is assumed that for any channel configuration which exceeds 400 channels of each type these results are sufficient to establish that simulation via the SDE or PCPA method are numerically equivalent (in terms of ISI statistics). Validation of the SDE model is important because all simulation methods outlined thus far have runtimes which scale with the number of channels, and significantly limit simulation of large (physiologically relevant) configurations. Since the runtime of the SDE model is independent of the configuration, simulation of the system defined in section 2.1.2 can now be openly, and quickly obtained across remainder of the channel number plane.

With the newly validated SDE model in hand, a survey of the ISI mean, variance, and CV surfaces are offered in Figure 5.6, which gives a broad survey up to 10,000

channels, and Figure 5.7, which gives a more detailed survey up to 3000 channels.

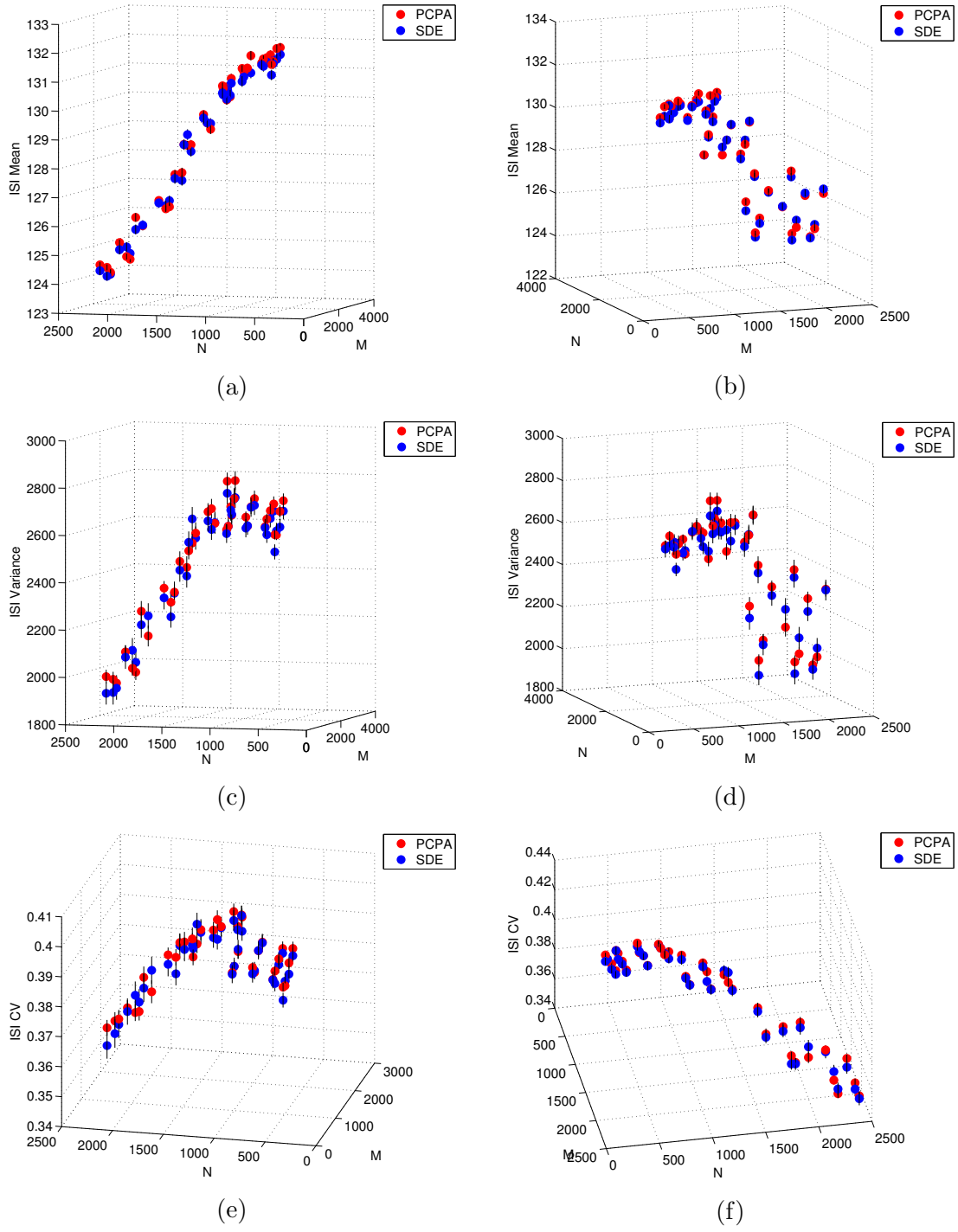


Figure 5.4: To establish convergence for various channel configurations, the system defined in section 2.1.2 is simulated using the PCPA (red, simulation method in section 4.2) and compared with the approximated system defined in section 5.1 (blue, simulation method in section 5.2). The ISI distribution of each channel configuration from each simulation method is found from the associated voltage trace (using a dual-threshold crossing method as described in section 2.1.3) and enables calculation of the ISI mean (a,b), variance (c,d), and CV (e,f). The standard error of each statistical measure is represented by black lines (often to small to be visible) and is obtained using approximations outlined in section 7.2. Note that each column offers a different view of the same figure. A close examination of all three statistical measures show that the PCPA and SDE models are within the standard error of each measure for channel configurations between 400 and 2400. Since the approximation is quantitatively validated in Figure 5.5 along the $N_N = N_M$ line, and since it should only get better as the channel numbers increase (see equation 5.8), this project makes the assumption that the PCPA and SDE models are numerically equivalent in terms of ISI statistics for all channel configurations containing more than 400 channels of each type.

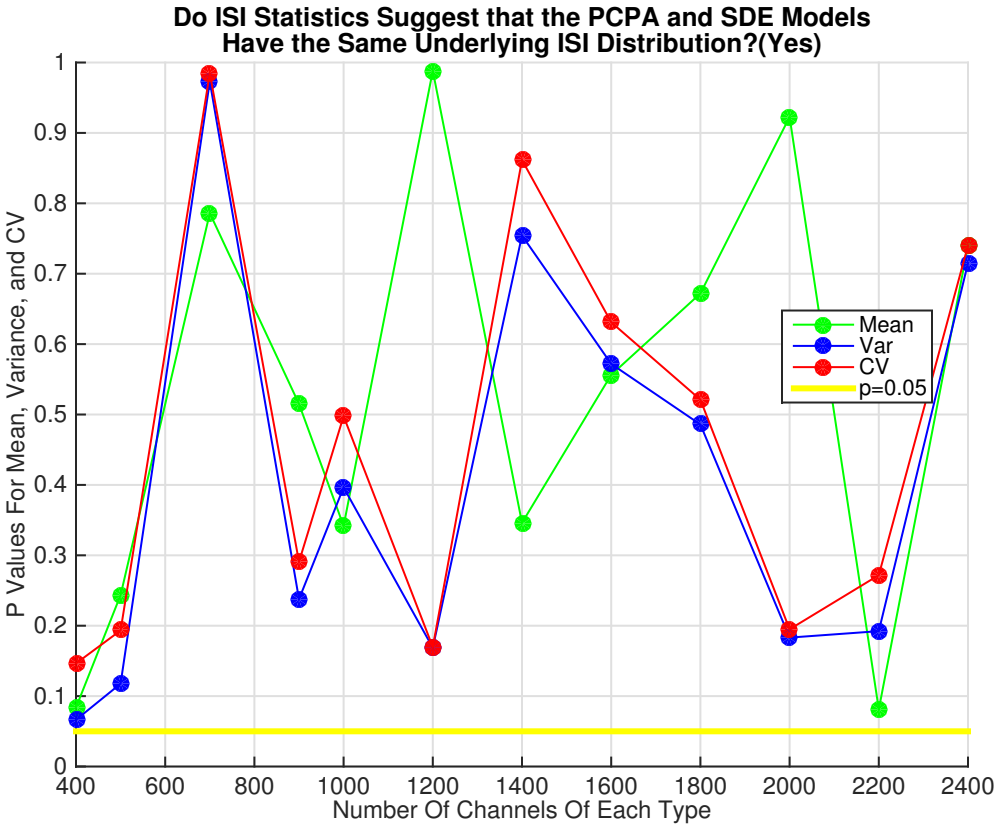


Figure 5.5: If N is the number of channels of each type, then under the null hypothesis that $H_0 : \mathrm{ISI}_N^{\mathrm{PCPA}} = \mathrm{ISI}_N^{\mathrm{SDE}}$, this plot shows the probability of observing values equal to or more extreme than the measured difference between the ISI mean, variance, and CV which result from the PCPA and SDE models (as depicted in Figure 5.4). The yellow line gives the standard significance level of 0.05, and clarifies that the measured ISI statistics do provide enough evidence to support the null hypothesis H_0. We can conclude that the SDE model converges to the PCPA model in a statistically significant way after 400 channels of each type are used. Therefore, the SDE model is considered valid for channel configurations along the $N_M = N_N$ line which exceed 400, as shown in Figure 1.1 which summarizes the results of this project. The hypothesis test was developed in section 7.4 and is given in equation 7.57 which uses the test statistic from equation 7.58. See figure 4.5 for an example of an unsuccessful test.

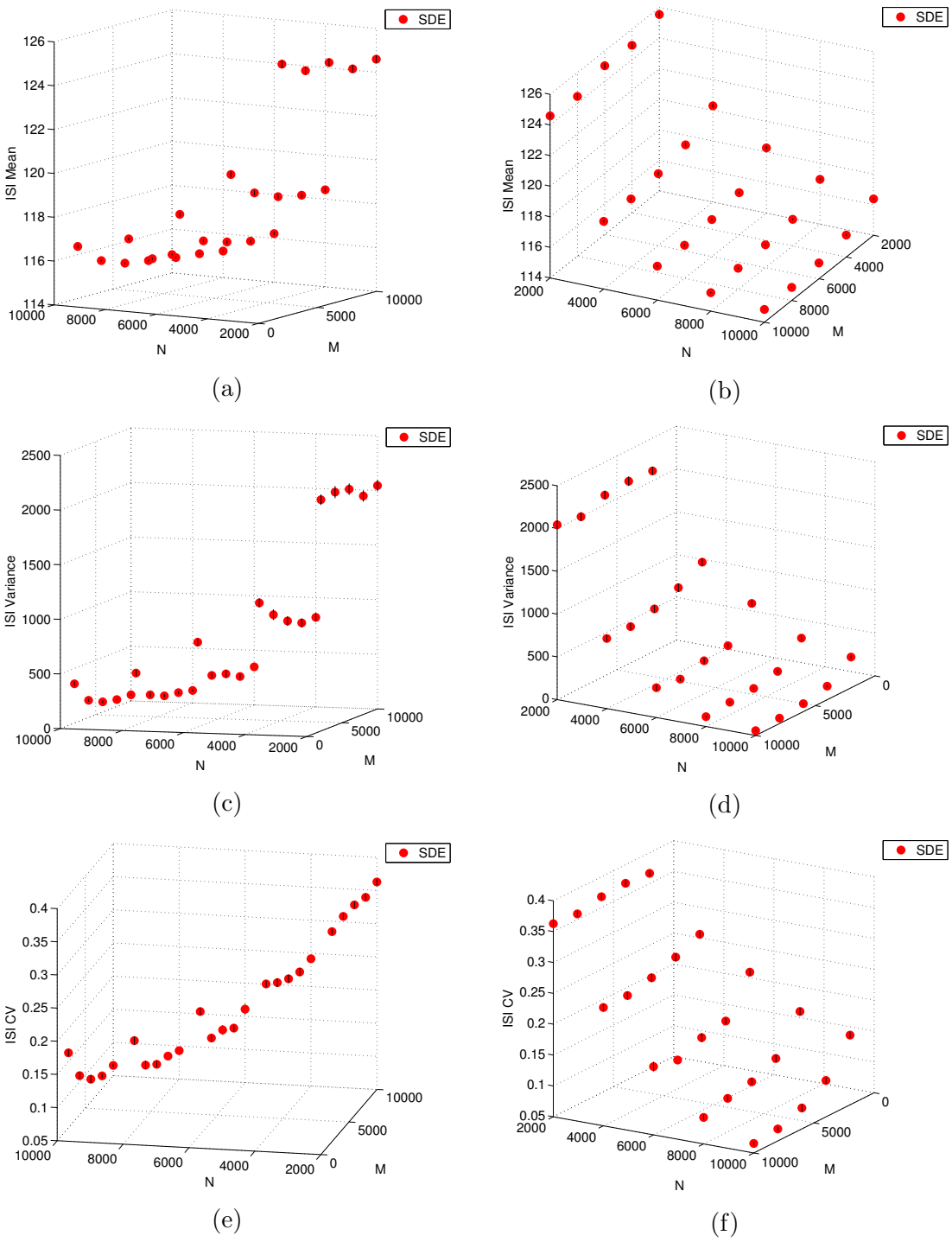


Figure 5.6: Now that the convergence of the SDE model to the PCPA model has been verified in Figure 5.4 and 5.5, it is assumed that any SDE simulation using a channel configuration exceeding 400 channels of each type is valid. This figure along with Figure 5.7, which offers a closer perspective, give an overview of the ISI mean, variance, and CV surfaces generated by the SDE model. The surface was generated by first simulating the SDE model defined in 5.1 using the algorithm defined in section 5.2. Once the voltage traces are simulated, the spike times are found using a dual-threshold crossing method (as described in section 2.1.3), and then the ISI mean (a,b), variance (c,d), and CV (e,f) are found from the resulting ISI distribution. The standard error of each statistical measure is represented by black lines (often to small to be visible) and is obtained using approximations outlined in section 7.2. Note that each column offers a different view of the same figure.

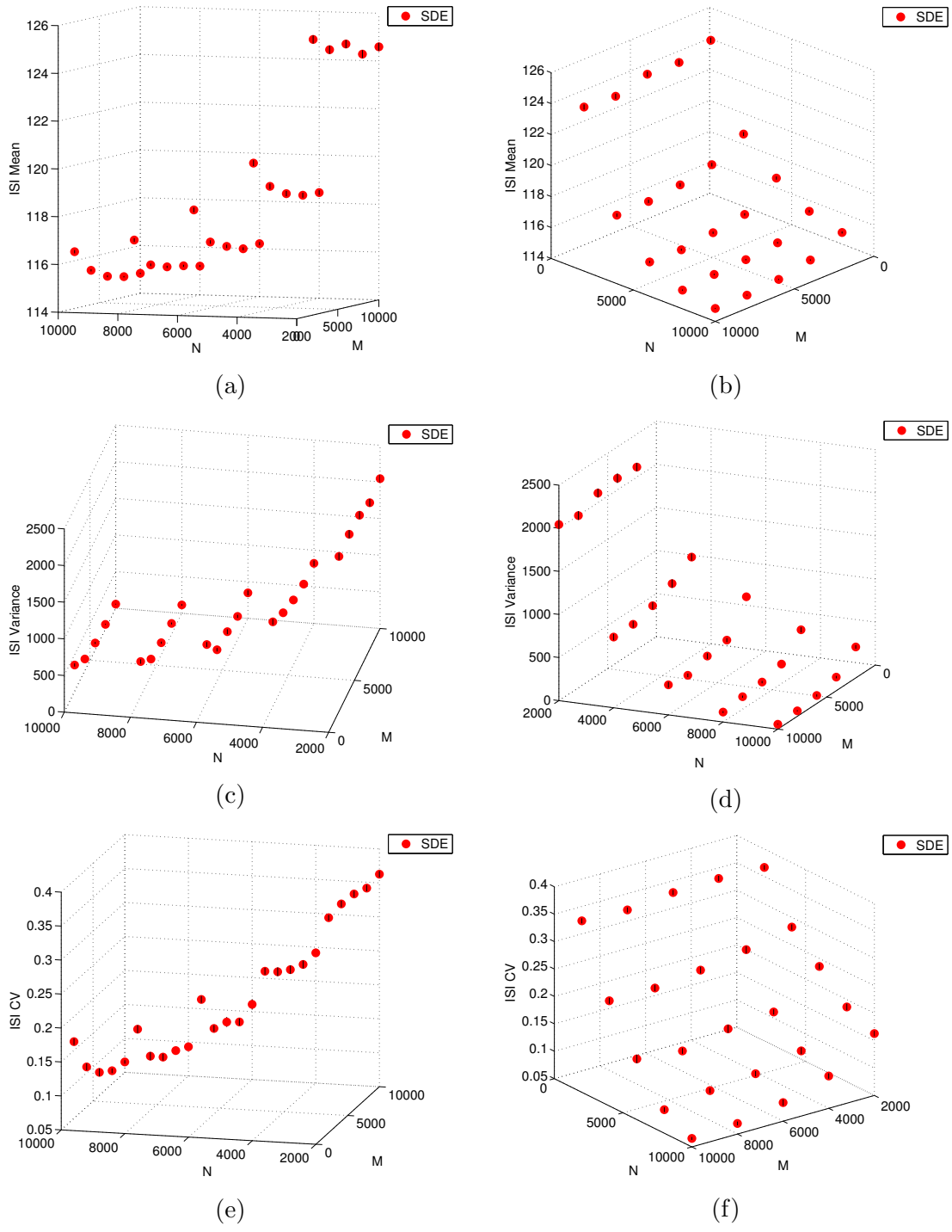


Figure 5.7: A closer perspective of the surface presented in Figure 5.6. Now that the convergence of the SDE model to the PCPA model has been verified in Figure 5.4 and 5.5, it is assumed that any SDE simulation using a channel configuration exceeding 400 channels of each type is valid. This figure along with Figure 5.6, which offers a broader perspective, give an overview of the ISI mean, variance, and CV surfaces generated by the SDE model. The surface was generated by first simulating the SDE model defined in 5.1 using the algorithm defined in section 5.2. Once the voltage traces are simulated, the spike times are found using a dual-threshold crossing method (as described in section 2.1.3), and then the ISI mean (a,b), variance (c,d), and CV (e,f) are found from the resulting ISI distribution. The standard error of each statistical measure is represented by black lines (often to small to be visible) and is obtained using approximations outlined in section 7.2. Note that each column offers a different view of the same figure.

Chapter 6

The Weak Noise Region

Ermentrout et al. applied weak Gaussian white noise to the voltage equation of a general, unspecified neural model which had an orbitally stable limit cycle solution, and derived an expression for the associated ISI variance of the model to first order [17]. The methods of this chapter extend the analysis in [17] through two critical observations: the methods of [17] also work with the application of Gaussian white noise on channel equations of this type of neural model, and the SDE model derived in chapter 5 is simply a deterministic system subject to scaled Gaussian white noise applied to the channel equation. The efforts of [17], these two observations, and a little bit of work will show that for sufficiently large channel numbers, the effect of channel noise on the ISI of the model derived in chapter 5 becomes so small that it can be linearized about the deterministic limit cycle solution with high accuracy, and related in closed form to the ISI variance (to first order).

After a short introductory section, this chapter derives a final result in two different ways. The first, in section 6.1, follows the approach of [17] to derive the following result,

$$\begin{aligned} \mathrm{Var}\left[\mathrm{ISI}\right] = \epsilon_{\mathrm{X}}^2 \int_0^T Z_{\mathrm{X}}^2(t)\ \sigma_{\mathrm{X}}^2(\gamma(t))dt \\ + \epsilon_{\mathrm{Y}}^2 \int_0^T Z_{\mathrm{Y}}^2(t)\ \sigma_{\mathrm{Y}}^2(\gamma(t))dt + O\left(\epsilon^2\right), \end{aligned} \tag{6.1}$$

which gives the variance of the ISI in terms of an integral around the deterministic limit cycle solution ($\gamma(t)$), with an integrand that includes the number of channels ($\epsilon_{\mathrm{X}}^2 = 1/N_{\mathrm{X}}$), the diffusion function from the SDE of each channel ($\sigma(t)$), and a function which gives the sensitivity of the system's period to state perturbations ($Z(t)$, the so-called infinitesimal phase response curve for the limit cycle solution). Afterwards, a second derivation is given in a detailed, step-by-step manner for the sake of pedological completeness and intuition. The two approaches meet at equation 6.84.

Finally, a comparison between the weak noise expression and the SDE model at the end of this chapter shows that each method produces ISI variance values which are distinct, but very close. The source of error most likely to arises from second order terms which appear in the weak noise expression, but that were ignored since they are nontrivial to calculate numerically.

6.0.1 Introduction to Neural Oscillators and Phase

Consider a system of autonomous ordinary differential equations in $\mathbb{R}^n$,

$$\frac{d\vec{X}}{dt} = F(\vec{X}), \tag{6.2}$$

which has a stable limit cycle solution, $\gamma(t)$, with period T. In our particular case,

$$\vec{X}(t) = \begin{bmatrix} V(t) \\ M(t) \\ N(t) \end{bmatrix}, \tag{6.3}$$

but for the sake of generality the presentation will be in n dimensions for as long as it is convenient. Note that the first row of the vector $\vec{X}$ will always hold the voltage, and the remaining dimensions will hold the channel state of the system. The T-periodic

limit cycle solution, $\gamma(t)$, has a phase which is defined as a function,

$$\theta : \gamma \to S^1 = [0, T), \tag{6.4}$$

where,

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1. \tag{6.5}$$

Though the choice of where $\theta(\gamma(t)) = 0$ is arbitrary, since neural oscillators periodically generate action potentials, by convention the point where the system reaches a maximum voltage is often considered to be zero phase and the next spike is phase T.

6.1 The Itō Approach to Weak Stochastic Perturbations

This section will follow [17] to succinctly derive equation 6.84. For a slower, more detailed approach see the next section. It was established in section 5.1 that for sufficiently large channel numbers the stochastic Morris Lecar model may be expressed in the form of an SDE,

$$\begin{bmatrix} dV \\ dX \\ dY \end{bmatrix} = \begin{bmatrix} F_{\mathrm{V}}(V,X,Y) \\ F_{\mathrm{X}}(V,X,Y) \\ F_{\mathrm{Y}}(V,X,Y) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \epsilon_{\mathrm{X}}\sigma_{\mathrm{X}}(V,X,Y)dW_{\mathrm{X}}(t) \\ \epsilon_{\mathrm{Y}}\sigma_{\mathrm{Y}}(V,X,Y)dW_{\mathrm{Y}}(t) \end{bmatrix}, \tag{6.6}$$

whose solution (when ϵ_{X} and ϵ_{Y} are zero) is a stable, T-periodic limit cycle $\gamma(\theta(t))$. Following the format of page 96 in [25], we can re-express this as,

$$d\vec{\mathbf{x}} = F(\vec{\mathbf{x}})dt + P(\vec{\mathbf{x}})dW, \tag{6.7}$$

where,

$$\vec{\mathbf{x}} = \begin{bmatrix} V \\ X \\ Y \end{bmatrix}, \quad P(\vec{\mathbf{x}}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \epsilon_{\mathrm{X}}\sigma_{\mathrm{X}}(\vec{\mathbf{x}}) & 0 \\ 0 & 0 & \epsilon_{\mathrm{Y}}\sigma_{\mathrm{Y}}(\vec{\mathbf{x}}) \end{bmatrix}, \quad dW = \begin{bmatrix} dW_{\mathrm{V}} \\ dW_{\mathrm{M}} \\ dW_{\mathrm{N}} \end{bmatrix}. \tag{6.8}$$

The application of the multi-dimensional Itō formula to $\theta(\vec{\mathbf{x}})$, as prescribed by [25], gives,

$$\begin{aligned} d\theta = &\left[F(\vec{\mathbf{x}}) \cdot \nabla\theta(\vec{\mathbf{x}}) + \frac{1}{2}\left(\epsilon_{\mathrm{X}}^2\sigma_{\mathrm{X}}^2(\vec{\mathbf{x}})\frac{\partial^2}{\partial X^2} + \epsilon_{\mathrm{Y}}^2\sigma_{\mathrm{Y}}^2(\vec{\mathbf{x}})\frac{\partial^2}{\partial Y^2}\right)\theta(\vec{\mathbf{x}})\right]dt \\ &+\left[\nabla(\theta(\vec{\mathbf{x}}))P(\vec{\mathbf{x}})\right]dW, \end{aligned} \tag{6.9}$$

Following [17], the second order terms (ϵ_{X}^2 and ϵ_{Y}^2) are neglected since second derivatives of θ are not easily attainable with numerical methods. However, in the small noise regime the gradient of θ is equivalent to a common function called the infinitesimal phase response curve, $\vec{Z}$, which gives the sensitivity of the system's period to small state perturbations, and is readily attainable through the solution of a linear differential equation called the adjoint equation [16]. The absence of these second order terms throughout the chapter are likely the cause of the discrepancy seen in the end result depicted in Figure 6.1. We are left with,

$$d\theta = F(\vec{\mathbf{x}}) \cdot \nabla\theta(\vec{\mathbf{x}})dt + \nabla\theta(\vec{\mathbf{x}})P(\vec{\mathbf{x}})dW. \tag{6.10}$$

As is seen in [17], $\theta(t) = t + O(\epsilon^2)$, so $\nabla\theta = \vec{Z}$ gives,

$$\begin{aligned}
\text{ISI} &= \theta(T) - \theta(0) \\
&= \int_0^T d\theta(\gamma(t)) \\
&= \int F(\gamma(t)) \cdot \nabla\theta(\gamma(t)) dt \\
&+ \epsilon_{\text{X}} \int_0^T \sigma_{\text{X}}(\gamma(t)) Z_{\text{X}}(\gamma(t)) dW_{\text{X}} \\
&+ \epsilon_{\text{Y}} \int_0^T \sigma_{\text{Y}}(\gamma(t)) Z_{\text{Y}}(\gamma(t)) dW_{\text{Y}} \\
&= T \\
&+ \epsilon_{\text{X}} \int_0^T \sigma_{\text{X}}(\gamma(t)) Z_{\text{X}}(\gamma(t)) dW_{\text{X}} \\
&+ \epsilon_{\text{Y}} \int_0^T \sigma_{\text{Y}}(\gamma(t)) Z_{\text{Y}}(\gamma(t)) dW_{\text{Y}}
\end{aligned} \tag{6.11}$$

where it is assumed that the noise amplitude, ϵ, is so small that the state of the system is always well approximated by the deterministic limit cycle solution, $\gamma(t)$. By assuming that noise sources are independent, and using Itō's isometry, we arrive at the final result,

$$\begin{aligned}
\text{Var}\left[\text{ISI}\right] = \epsilon_{\text{X}}^2 &\int_0^T Z_{\text{X}}^2(t)\ \sigma_{\text{X}}^2(\gamma(t)) dt \\
+ \epsilon_{\text{Y}}^2 &\int_0^T Z_{\text{Y}}^2(t)\ \sigma_{\text{Y}}^2(\gamma(t)) dt + O\left(\epsilon^2\right).
\end{aligned} \tag{6.12}$$

Now that the result has been derived using Itō's formula following [17], we proceed with an alternative derivation from first principles for the sake of a full pedological presentation and to develop some intuition. The following derivation reaches the final result shown above at equation 6.84 on page 111 (a savvy reader may wish to move there immediately).

6.2 Weak Deterministic Perturbations of Neural Oscillators

A direct relationship between ISI variability and weak channel noise will be derived in two large steps which each contain three small steps. Firstly, the relationship between deterministic perturbations and spike timing is described using a simple theory of weak singly perturbed neural oscillators, which is then expanded to a theory of neural oscillators subject to N weak perturbations, and finally extended (through a limit) to neural oscillators subject to weak continuous perturbations. Secondly, the relationship between stochastic perturbations and spike timing is described by repeating the same sequence of steps (single, multiple, and then continuous perturbations), but with stochastic perturbations. A simple manipulation of the effect of stochastic perturbations on spike timing will yield the ISI distribution, to first order, of the neural model subject to stochastic perturbations.

Though it may be natural to think of neurons (and stochastic neural models) as having an intrinsic source of channel noise which is somehow inseparable from its deterministic dynamics, a slight change in perspective enables us to use the tools described in the last paragraph to great effect; Stochastic neural models (and perhaps neurons) can be thought of as a deterministic system subject to stochastic perturbations. The third step of this chapter will be to cast channel noise of the Morris-Lecar model defined in chapter 5 as a series of weak, continuous, stochastic perturbations which act upon the deterministic Morris-Lecar model.

After stochastic Morris-Lecar perturbations are defined, the ISI distribution of the deterministic system under the influence of such perturbations can be derived in closed form, where each part of the expression is numerically obtainable. The last section of the chapter uses these observations to numerically calculate the ISI mean and variance which results from a weakly perturbed, stochastic Morris-Lecar

neuron, and compares the same quantities which result from the SDE model. Since the ISI mean and variance can be calculated almost instantly from the weak noise expression, this chapter is a final step in this project which sought to find fast methods for simulating the stochastic Morris-Lecar model originally defined in section 2.1.2.

6.2.1 Weak Individual Perturbations

If, during the spontaneous firing of an oscillator defined by equation 6.2, a small current I_0 is injected for a short time ω, then a charge $I_0\omega$ is injected into the cell and the voltage is increased by $\beta_{\mathrm{V}} = I_0\omega/C$. Such a voltage perturbation at a time $0 \leq \tau < T$ can be represented in the dynamic equation as,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \delta(t-\tau)\epsilon\vec{\beta}, \tag{6.13}$$

where $\delta(t-\tau)$ is the Dirac delta function, and,

$$\vec{\beta} = \begin{bmatrix} \beta_{\mathrm{V}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{6.14}$$

so the additional term is understood (for now) to act exclusively on the voltage equation. Given that ϵ is sufficiently small, we may retain phase coordinates by augmenting equation 6.5 with a term that represents the response of the oscillator to the perturbation in terms of phase [32, 34],

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \vec{Z}(\theta)\cdot\delta(t-\tau)\epsilon\vec{\beta} + o(\epsilon). \tag{6.15}$$

Central to the new term is the infinitesimal phase response curve (iPRC), $\vec{Z}(\theta)$, which is in units of phase per millivolt. The iPRC is derived from a more intuitive function called the phase response curve (PRC) which measures the effect of a state perturbation on the timing of the next spike. The PRC is defined as,

$$R(\vec{\beta}, \tau) = T - \hat{T}(\vec{\beta}, \tau), \tag{6.16}$$

where T is the period of the system governed by equation 6.2, $\hat{T}(\vec{\beta}, \tau)$ is the period of the system governed by equation 6.13, and τ is the time of the perturbation. Given T and θ (as defined by equation 6.15), we can also define $\hat{T}$ up to first order as the solution, t_s, to $\theta(t_s) = T$. The iPRC comes from expanding R (we assume the function is smooth throughout this chapter) about $\vec{\beta} = 0$ and then evaluating it for some small perturbation $\epsilon\vec{\beta}$,

$$R(\epsilon\vec{\beta}, \tau) = \nabla_\beta R(0, \tau) \cdot \epsilon\vec{\beta} + O\left(\epsilon^2\right), \tag{6.17}$$

where the gradient term is the iPRC used in equation 6.15,

$$\vec{Z}(\theta(\tau)) = \nabla_\beta R(0, \tau), \tag{6.18}$$

and is understood to be the the gradient of the PRC evaluated around the limit cycle solution of equation 6.2. Given a system like equation 6.2, the iPRC can be found by numerically solving a linear differential equation, which arises in any theory of small perturbations of limit cycles, called the adjoint equation [16],

$$\frac{d}{dt}\vec{Z}(t) = -DF^T(\gamma(t))\vec{Z}(t). \tag{6.19}$$

Note that though the the perturbation vector was defined as,

$$\vec{\beta} = \begin{bmatrix} \beta_{\mathrm{V}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{6.20}$$

in general the perturbation can act on the channel states as well. In our specific case, it would look like this,

$$\vec{\beta} = \begin{bmatrix} \beta_{\mathrm{V}} \\ \beta_{\mathrm{M}} \\ \beta_{\mathrm{N}} \end{bmatrix}, \tag{6.21}$$

where $\beta_{\mathrm{V}}, \beta_{\mathrm{M}}$, and β_{N} are the perturbations of the voltage, M channel, and N channel state variables, respectively. Making small perturbations to channel states is really only sensible in systems which describe channel states using continuous variables (like the standard Hodgkin-Huxley model), or systems which are large enough to be well approximated by a continuous channel state model (just as the model defined in section 2.1.2 was approximated by the model in section 5.1).

In this section a system of autonomous ordinary differential equations with a stable, T-periodic limit cycle solution was defined. A perturbation was introduced into the system and it was established that, for sufficiently small perturbations, phase coordinates could be retained and that the sensitivity of the oscillator to small perturbations is described by the iPRC which defines a linear relationship between the perturbation and the resulting alteration of the subsequent action potential. We now understand the simplest form of an idea which acts as the central tool of this section: small perturbations which act on state have a linear affect on the timing of the next spike. A progressive generalization of this concept will lead to the final result of the chapter. In the next section we will extend these ideas to systems which are subject

to many perturbations.

6.2.2 Weak Finite Perturbations

In this section, a generalization of equation 6.17 will be derived which represents the phase response curve (PRC) of a neural oscillator subject to many perturbations. The dynamic equation is as follows,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon \sum_{k=1}^{N} \delta(t - \tau_k)\vec{\beta}_{\tau_k}, \tag{6.22}$$

where $0 \leq \tau_1 < \tau_2 < ... < \tau_N < T$, and $\delta(t - \tau_k)$ is the Dirac delta function. We can then imagine a multidimensional PRC expressed as,

$$R\left(\vec{\beta_1}, ..., \vec{\beta_N}, \tau_1, ..., \tau_N\right) = T - \hat{T}\left(\vec{\beta_1}, ..., \vec{\beta_N}, \tau_1, ..., \tau_N\right), \tag{6.23}$$

or in our particular case as,

$$R\left(\epsilon\vec{\beta_1}, ..., \epsilon\vec{\beta_N}, \tau_1, ..., \tau_N\right) = T - \hat{T}\left(\epsilon\vec{\beta_1}, ..., \epsilon\vec{\beta_N}, \tau_1, ..., \tau_N\right). \tag{6.24}$$

This potentially complicated expression can be approximated to first order for sufficiently small ϵ about $\vec{\beta_k} = 0$ for all k,

$$R\left(\epsilon\vec{\beta_1}, ..., \epsilon\vec{\beta_N}, \tau_1, ..., \tau_N\right) = \epsilon \sum_{k=1}^{N} \nabla_\beta R(0, \theta(\tau_k)) \cdot \vec{\beta}_{\tau_k} + O\left(\epsilon^2\right), \tag{6.25}$$

or using the notation for the iPRC,

$$R\left(\epsilon\vec{\beta_1}, ..., \epsilon\vec{\beta_N}, \tau_1, ..., \tau_N\right) = \epsilon \sum_{k=1}^{N} \vec{Z}(\theta(\tau_k)) \cdot \vec{\beta}_{\tau_k} + O\left(\epsilon^2\right), \tag{6.26}$$

where $\theta(t)$ is given by,

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \epsilon \sum_{k=1}^{N} \delta(t-\tau_k)\vec{Z}(\theta(t)) \cdot \vec{\beta}_{\tau_k} + o(\epsilon). \tag{6.27}$$

Since the system defined in equation 6.2 was assumed to have a stable limit cycle, we expect perturbations away from the solution to correct, but to advance or delay the phase. So, while t is a measure of "lab time", $\theta(t)$ can be thought of as "neuron time" which describes how far along the system is in its cycle. Perturbations cause these two measures to go out of sync, which is why equation 6.26 is expressed with $\vec{Z}(\theta(t))$ (the sensitivity of the system to perturbations as a function of "neuron time") as opposed to $\vec{Z}(t)$ (the sensitivity of the system to perturbations as a function of "lab time"). Equation 6.26 can be interpreted as follows: as long the preceding perturbations leave the solution close to the limit cycle solution, the current perturbation can be considered to act on an unperturbed system (as in equation 6.17) in "neuron time" given by θ.

6.2.3 Weak Continuous Perturbations

In this section we will take a many perturbation limit to get expressions in terms of continuous perturbation functions. Recall the dynamic equation in the last section,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon \sum_{k=1}^{N} \delta(t-\tau_k)\vec{\beta}_{\tau_k}. \tag{6.28}$$

Moving forward requires a slight change in perspective. Imagine that the number of perturbations increases ($N \to \infty$) or that each perturbation occurs over a longer period of time so that the perturbations can be naturally modeled by a perturbation

function. In our specific case it would appear as follows,

$$\vec{P}(t) = \begin{bmatrix} \beta_{\mathrm{V}}(t) \\ \beta_{\mathrm{M}}(t) \\ \beta_{\mathrm{N}}(t) \end{bmatrix}. \tag{6.29}$$

The new dynamic equation would then be,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon\vec{P}(t). \tag{6.30}$$

This new system gives a version of the PRC which is analogous to equation 6.23,

$$R\left[\epsilon\vec{P}(t)\right] = T - \hat{T}\left[\epsilon\vec{P}(t)\right], \tag{6.31}$$

where now both R and $\hat{T}$ are functionals. For sufficiently small ϵ, we can approximate this function with equation 6.25 by sampling $\epsilon\vec{P}(t)$ at times $[t_1, ..., t_N]$, where for all k where $1 \leq k \leq N$, $t_k < T$, and $t_{k+1} - t_k = \Delta t = T/N$,

$$R\left(\epsilon\vec{P}(t_1), ..., \epsilon\vec{P}(t_N), t_1, ..., t_N\right) = \epsilon\sum_{k=1}^{N} \nabla_\beta R(0, \theta(t_k)) \cdot \vec{P}(t_k)\Delta t + O\left(\epsilon^2\right). \tag{6.32}$$

In the limit where, $N \to \infty$, the approximation becomes exact[1],

$$\begin{aligned} R\left[\epsilon\vec{P}(t)\right] &= \lim_{N\to\infty} R\Big(\epsilon\vec{P}(t_1), ..., \epsilon\vec{P}(t_N), t_1, ..., t_N\Big) \\ &= \lim_{N\to\infty} \epsilon\sum_{k=1}^{N} \nabla_\beta R(0, \theta(t_k)) \cdot \vec{P}(t_k)\Delta t + O\left(\epsilon^2\right) \\ &= \epsilon\int_0^T \nabla_\beta R(0, \theta(t)) \cdot \vec{P}(t)dt + O\left(\epsilon^2\right), \end{aligned} \tag{6.33}$$

[1] There are a number of technical conditions in this section and the analogous section on continuous stochastic perturbations that are being overlooked, partially because they are beyond the scope of this project, and partially because the practical application of these claims would probably satisfy the conditions anyway. In any case, it is acknowledge that formal proof requires more rigor, so the project proceeds with the assumption that the necessary conditions hold, and will indirectly validate its claims with the aid of a numerical test included at the end of the chapter.

or in terms of the iPRC,

$$R\left[\epsilon\vec{P}(t)\right] = \epsilon\int_0^T \vec{Z}(\theta(t))\cdot\vec{P}(t)dt + O\left(\epsilon^2\right), \tag{6.34}$$

where $\theta(t)$ is given by,

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \epsilon\,\vec{Z}(\theta(t))\cdot\vec{P}(t) + o(\epsilon). \tag{6.35}$$

Intuitively, equation 6.34 says that the effect of perturbations on the spike timing add, and that the effect of each source of perturbation on the spike timing is simply the magnitude of that perturbation times the sensitivity of the system to the perturbation at the phase (or "neuron time") the perturbation is received (up to first order in ϵ). Note that in the functional case, $\vec{Z}(\theta(t))$ acts like the functional derivative of the PRC evaluated at a function $f(x) = 0$.

The function $\vec{Z}(\theta(t))$ can be found numerically by solving equation 6.19, and that in many cases (like one which will be presented soon) $\vec{P}(t)$ is given. Under these conditions, $\theta(t)$ is the only unknown. However, since there is already a pre-factor of ϵ in equations 6.34, a short calculation shows that $\theta(t)$ can be replaced by its zeroth order approximation, t. If $\theta(t)$ is expanded in powers of ϵ

$$\theta(t) = t + O\left(\epsilon\right), \tag{6.36}$$

and substituted into $\vec{Z}(\theta)$,

$$\vec{Z}(t + O\left(\epsilon\right)) = \vec{Z}(t) + O\left(\epsilon\right), \tag{6.37}$$

it is clear that equation 6.34 can be re-written to first order as,

$$R\left[\epsilon\vec{P}(t)\right] = \epsilon\int_0^T \vec{Z}(t)\ \cdot\vec{P}(t)dt + O\left(\epsilon^2\right). \tag{6.38}$$

6.3 Weak Stochastic Perturbations of Neural Oscillators

The progressive generalization of the single perturbation PRC to the continuously perturbed PRC performed in the last section will be redone in the this section with stochastic perturbations. Starting from a single stochastic perturbation, we generalize to a finite number of stochastic perturbations, and then take a limit to get an expression for the PRC subject to weak continuous stochastic perturbations. The last section will define a weak continuous stochastic perturbation function called the stochastic Morris-Lecar perturbations, and will implement them in an expression for the PRC to eventually yield the ISI mean and variance of the weak-noise Morris-Lecar neuron.

6.3.1 Weak Stochastic Individual Perturbations

Recall the dynamic equation presented in section 6.2.1,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \delta(t-\tau)\epsilon\vec{\beta}, \tag{6.39}$$

where the first component of $\vec{\beta}$ is voltage and the remaining components describe the channel state of the system. In our specific case (from now on we will derive the results using the specific 3-dimensional system of this project for notational convenience, even

though the results apply to higher dimensional systems),

$$\vec{\beta} = \begin{bmatrix} \beta_{\mathrm{V}} \\ \beta_{\mathrm{M}} \\ \beta_{\mathrm{N}} \end{bmatrix}. \tag{6.40}$$

We can make the perturbation random by introducing three mutually independent Gaussian random variables,

$$\begin{aligned} \xi_{\mathrm{V}} &\sim \mathcal{N}(0,1), \\ \xi_{\mathrm{M}} &\sim \mathcal{N}(0,1), \\ \xi_{\mathrm{N}} &\sim \mathcal{N}(0,1), \end{aligned} \tag{6.41}$$

and forcing $0 \leq \beta_{\mathrm{V}}, \beta_{\mathrm{M}}, \beta_{\mathrm{N}}$ so that,

$$\vec{\beta}^{*} = \begin{bmatrix} \beta_{\mathrm{V}}\xi_{\mathrm{V}} \\ \beta_{\mathrm{M}}\xi_{\mathrm{M}} \\ \beta_{\mathrm{N}}\xi_{\mathrm{N}} \end{bmatrix}. \tag{6.42}$$

The dynamic equation subject to one stochastic perturbation at time τ is then,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \delta(t-\tau)\epsilon\vec{\beta}^{*}. \tag{6.43}$$

Proceeding as we did in section 6.2.1, given that ϵ is sufficiently small, we may retain phase coordinates by augmenting equation 6.5 with a term that represents the response of the oscillator to the perturbation in terms of phase [32, 34],

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \vec{Z}(\theta)\cdot\delta(t-\tau)\epsilon\vec{\beta}^{*} + o(\epsilon). \tag{6.44}$$

The PRC subject to a single random perturbation is then,

$$R(\vec{\beta}^*, \tau) = T - \hat{T}(\vec{\beta}^*, \tau), \tag{6.45}$$

where $R(\vec{\beta}^*, \tau)$ is now a distribution. By expanding around $\vec{\beta} = 0$,

$$R(\epsilon\vec{\beta}^*, \tau) = \nabla_\beta R(0, \tau) \cdot \epsilon\vec{\beta}^* + O\left(\epsilon^2\right), \tag{6.46}$$

or

$$R(\epsilon\vec{\beta}^*, \tau) = \vec{Z}(\theta(t)) \cdot \epsilon\vec{\beta}^* + O\left(\epsilon^2\right), \tag{6.47}$$

we can see that to first order $R(\vec{\beta}^*, \tau)$ is a simply a Gaussian distribution with,

$$R(\vec{\beta}^*, \tau) \sim \epsilon\mathcal{N}\left(0, \sum_{K=\mathrm{V,M,N}} \left(\frac{\partial R(0,\tau)}{\partial \mathrm{K}}\beta_K\right)^2\right) + O\left(\epsilon^2\right). \tag{6.48}$$

Note that the variance of the expression above is to second order, but that there are also unspecified terms of second order within the $O\left(\epsilon^2\right)$ term. This problem persists as we generalize further, and is likely the source of error which is seen in the figure at the end of the chapter. The alternative is to expand to second order in equation 6.46, but this leads to second derivatives of the phase response term, which are not readily obtainable.

6.3.2 Weak Stochastic Finite Perturbations

Next we will generalize to a finite number of stochastic perturbations. Consider the dynamic equation which appeared in section 6.2.2,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon\sum_{k=1}^{N} \delta(t - \tau_k)\vec{\beta}_{\tau_k}, \tag{6.49}$$

where $0 \leq \tau_1 < \tau_2 < ... < \tau_N < \tau_{N+1} \equiv T$, and $\delta(t - \tau_k)$ is the Dirac delta function. The perturbations are made random by introducing three sequences of mutually independent Gaussian random variables,

$$\begin{aligned} \left[\xi_{\mathrm{V}_{\tau_k}}\right]_{k=1}^{N} &\sim \mathcal{N}(0,1), \\ \left[\xi_{\mathrm{M}_{\tau_k}}\right]_{k=1}^{N} &\sim \mathcal{N}(0,1), \\ \left[\xi_{\mathrm{N}_{\tau_k}}\right]_{k=1}^{N} &\sim \mathcal{N}(0,1), \end{aligned} \tag{6.50}$$

where for each $Q = [\mathrm{V}, \mathrm{M}, \mathrm{N}]$, $\langle \xi_{Q_{\tau_k}} \xi_{Q'_{\tau_{k'}}} \rangle = \delta(Q_{\tau_k} - Q'_{\tau_{k'}})$, and by forcing $0 \leq \beta_{\mathrm{V}}, \beta_{\mathrm{M}}, \beta_{\mathrm{N}}$ so that,

$$\vec{\beta}^*_{\tau_k} = \begin{bmatrix} \beta_{\mathrm{V}_{\tau_k}} \xi_{\mathrm{V}_{\tau_k}}, \\ \beta_{\mathrm{M}_{\tau_k}} \xi_{\mathrm{M}_{\tau_k}}, \\ \beta_{\mathrm{N}_{\tau_k}} \xi_{\mathrm{N}_{\tau_k}}, \end{bmatrix}. \tag{6.51}$$

The dynamic equation then becomes,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon \sum_{k=1}^{N} \delta(t - \tau_k) \vec{\beta}^*_{\tau_k}. \tag{6.52}$$

The PRC in this case becomes,

$$R\left(\epsilon\vec{\beta}^*_1, ..., \epsilon\vec{\beta}^*_N, \tau_1, ..., \tau_N\right) = T - \hat{T}\left(\epsilon\vec{\beta}^*_1, ..., \epsilon\vec{\beta}^*_N, \tau_1, ..., \tau_N\right), \tag{6.53}$$

which can be approximated to first order for sufficiently small ϵ about $\vec{\beta}_k = 0$ for all k,

$$R\left(\epsilon\vec{\beta}^*_1, ..., \epsilon\vec{\beta}^*_N, \tau_1, ..., \tau_N\right) = \epsilon \sum_{k=1}^{N} \nabla_\beta R(0, \theta(\tau_k)) \cdot \vec{\beta}^*_{\tau_k} + O\left(\epsilon^2\right), \tag{6.54}$$

or using the notation for the iPRC,

$$R\left(\epsilon\vec{\beta}^*_1, ..., \epsilon\vec{\beta}^*_N, \tau_1, ..., \tau_N\right) = \epsilon \sum_{k=1}^{N} \vec{Z}(\theta(\tau_k)) \cdot \vec{\beta}^*_{\tau_k} + O\left(\epsilon^2\right), \tag{6.55}$$

where $\theta(t)$ is given by,

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \epsilon \sum_{k=1}^{N} \delta(t-\tau_k)\vec{Z}(\theta(t)) \cdot \vec{\beta}^*_{\tau_k} + o(\epsilon). \tag{6.56}$$

We can see both from equation 6.55 and by extension of equation 6.48 that,

$$\begin{aligned} R\Big(\epsilon\vec{\beta}^*_1, ...,&\epsilon\vec{\beta}^*_N, \tau_1, ..., \tau_N\Big) \sim \\ &\epsilon\,\mathcal{N}\left(0, \sum_{k=1}^{N} \sum_{Q=\mathrm{V,M,N}} \left(\frac{\partial R(0,\tau_k)}{\partial \mathrm{Q}} \beta_{Q_{\tau_k}}\right)^2\right) + O\left(\epsilon^2\right). \end{aligned} \tag{6.57}$$

Note that, just like equation 6.48, there are unspecified second order terms which may interact with the specified terms. We ignore these terms because they are not readily obtainable, even though they are likely the source of error in the numerical test included at the end of the chapter.

6.3.3 Weak Stochastic Continuous Perturbations

In this section we will take a many perturbation limit of the equations derived in the last section to get expressions in terms of stochastic continuous perturbation functions. Recall the dynamic equation in the last section,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon \sum_{k=1}^{N} \delta(t-\tau_k)\vec{\beta}^*_{\tau_k}. \tag{6.58}$$

Moving forward in the same fashion as section 6.2.3, imagine that the number of perturbations increases ($N \to \infty$) or that each perturbation occurs over a longer period of time so that the perturbations can be naturally modeled by a perturbation function,

$$\vec{P}(t) = \begin{bmatrix} \beta_{\mathrm{V}}(t) \\ \beta_{\mathrm{M}}(t) \\ \beta_{\mathrm{N}}(t) \end{bmatrix}. \tag{6.59}$$

The perturbations are made random by forcing $0 \leq \beta_{\mathrm{V}}(t), \beta_{\mathrm{M}}(t), \beta_{\mathrm{N}}(t)$ for all t, and introducing three mutually independent Gaussian white noise processes, $\xi_{\mathrm{V}}(t), \xi_{\mathrm{M}}(t), \xi_{\mathrm{N}}(t)$, where $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$ so that,

$$\vec{P}^*(t) = \begin{bmatrix} \beta_{\mathrm{V}}(t)\xi_{\mathrm{V}}(t) \\ \beta_{\mathrm{M}}(t)\xi_{\mathrm{M}}(t) \\ \beta_{\mathrm{N}}(t)\xi_{\mathrm{N}}(t) \end{bmatrix}. \tag{6.60}$$

The new dynamic equation then becomes,

$$\frac{d\vec{X}}{dt} = F(\vec{X}) + \epsilon \vec{P}^*(t). \tag{6.61}$$

This new system gives a version of the PRC which is analogous to equation 6.53,

$$R\left[\epsilon \vec{P}^*(t)\right] = T - \hat{T}\left[\epsilon \vec{P}^*(t)\right], \tag{6.62}$$

where now both R and $\hat{T}$ are functionals.

We can approximate this function with equation 6.54 by sampling $\epsilon \vec{P}^*(t)$ at times $[t_1, ..., t_N]$, where for all k, such that $1 \leq k \leq N$, $t_k < T$, and $t_{k+1} - t_k = \Delta t = T/N$,

$$R\left(\epsilon \vec{P}^*(t_1), ..., \epsilon \vec{P}^*(t_N), t_1, ..., t_N\right) = \epsilon \sum_{k=1}^{N} \nabla_\beta R(0, \theta(t_k)) \cdot \vec{P}^*(t_k)\Delta t + O\left(\epsilon^2\right). \tag{6.63}$$

In the limit where, $N \to \infty$, the approximation becomes exact[2],

$$\begin{aligned} R\left[\epsilon \vec{P}^*(t)\right] &= \lim_{N\to\infty} R\Big(\epsilon \vec{P}^*(t_1), ..., \epsilon \vec{P}^*(t_N), t_1, ..., t_N\Big) \\ &= \lim_{N\to\infty} \epsilon \sum_{k=1}^{N} \nabla_\beta R(0, \theta(t_k)) \cdot \vec{P}^*(t_k)\Delta t + O\left(\epsilon^2\right) \\ &= \epsilon \int_0^T \nabla_\beta R(0, \theta(t)) \cdot \vec{P}^*(t) dt + O\left(\epsilon^2\right), \end{aligned} \tag{6.64}$$

[2] See footnote 1 just above the analogous deterministic equation 6.33

or in terms of the iPRC,

$$R\left[\epsilon\vec{P}^*(t)\right] = \epsilon\int_0^T \vec{Z}(\theta(t))\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right), \tag{6.65}$$

where $\theta(t)$ is given by,

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \epsilon\,\vec{Z}(\theta(t))\cdot\vec{P}^*(t) + o(\epsilon). \tag{6.66}$$

Using the same argument that was used to derive equation 6.38, since $\theta(t)$ already has a pre-factor of ϵ we can substitute its zeroth order approximation in both equations above,

$$R\left[\epsilon\vec{P}^*(t)\right] = \epsilon\int_0^T \vec{Z}(t)\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right), \tag{6.67}$$

and,

$$\frac{d}{dt}\left(\,\theta(\,\gamma(t)\,)\,\right) = 1 + \epsilon\,\vec{Z}(t)\cdot\vec{P}^*(t) + o(\epsilon). \tag{6.68}$$

Evaluation of the dot product shows the independent contributions of each type of state perturbation to the PRC,

$$\begin{aligned} R\left[\epsilon\vec{P}^*(t)\right] &= \epsilon\int_0^T \vec{Z}(t)\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right) \\ &= \epsilon\int_0^T Z_{\mathrm{V}}(t)\beta_{\mathrm{V}}(t)\xi_{\mathrm{V}}(t)dt \\ &+ \epsilon\int_0^T Z_{\mathrm{M}}(t)\beta_{\mathrm{M}}(t)\xi_{\mathrm{M}}(t)dt \\ &+ \epsilon\int_0^T Z_{\mathrm{N}}(t)\beta_{\mathrm{N}}(t)\xi_{\mathrm{N}}(t)dt + O\left(\epsilon^2\right), \end{aligned} \tag{6.69}$$

where $Z_{\mathrm{V}}, Z_{\mathrm{M}}, Z_{\mathrm{N}}$ are the voltage, M channel, and N channel components of the iPRC, respectively. If $dW_{\mathrm{V}}, dW_{\mathrm{M}}, dW_{\mathrm{N}}$ are the Brownian increments associated with

$\xi_{\mathrm{V}}(t), \xi_{\mathrm{M}}(t), \xi_{\mathrm{N}}(t)$, then the equation above can be re-expressed as,

$$\begin{aligned} R\left[\epsilon\vec{P}^*(t)\right] &= \epsilon\int_0^T \vec{Z}(t)\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right) \\ &= \epsilon\int_0^T Z_{\mathrm{V}}(t)\beta_{\mathrm{V}}(t)dW_{\mathrm{V}} \\ &+ \epsilon\int_0^T Z_{\mathrm{M}}(t)\beta_{\mathrm{M}}(t)dW_{\mathrm{M}} \\ &+ \epsilon\int_0^T Z_{\mathrm{N}}(t)\beta_{\mathrm{N}}(t)dW_{\mathrm{N}} + O\left(\epsilon^2\right). \end{aligned} \tag{6.70}$$

By recalling the definition of the PRC detailed in equation 6.62, this expression can be used to write the interspike interval distribution of an oscillator defined by equation 6.61, and subjected to $\vec{P}^*(t)$,

$$\begin{aligned} \hat{T}\left[\epsilon\vec{P}^*(t)\right] &= T - \epsilon\int_0^T \vec{Z}(t)\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right) \\ &= T - \epsilon\int_0^T Z_{\mathrm{V}}(t)\beta_{\mathrm{V}}(t)dW_{\mathrm{V}} \\ &- \epsilon\int_0^T Z_{\mathrm{M}}(t)\beta_{\mathrm{M}}(t)dW_{\mathrm{M}} \\ &- \epsilon\int_0^T Z_{\mathrm{N}}(t)\beta_{\mathrm{N}}(t)dW_{\mathrm{N}} + O\left(\epsilon^2\right). \end{aligned} \tag{6.71}$$

Since the white noise is symmetric, it is now easy to see that,

$$\mathbf{E}\left[\hat{T}\left[\epsilon\vec{P}^*(t)\right]\right] = T + O\left(\epsilon^2\right). \tag{6.72}$$

The variance can also be obtained,

$$\begin{aligned} \mathrm{Var}\left[\hat{T}\left[\epsilon\vec{P}^*(t)\right]\right] &= \mathrm{Var}\left[T - \epsilon\int_0^T \vec{Z}(t)\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right)\right] \\ &= \mathrm{Var}\left[T\right] + \mathrm{Var}\left[\epsilon\int_0^T \vec{Z}(t)\cdot\vec{P}^*(t)dt + O\left(\epsilon^2\right)\right], \end{aligned} \tag{6.73}$$

assuming that all terms in $O\left(\epsilon^2\right)$ are independent of the stochastic integral, or at

least do not affect moments of order 2 or lower, and noting that the noise sources are independent yields,

$$\begin{aligned}
&= \mathrm{Var}\left[\epsilon \int_0^T Z_{\mathrm{V}}(t)\beta_{\mathrm{V}}(t)dW_{\mathrm{V}}\right] \\
&+ \mathrm{Var}\left[\epsilon \int_0^T Z_{\mathrm{M}}(t)\beta_{\mathrm{M}}(t)dW_{\mathrm{M}}\right] \\
&+ \mathrm{Var}\left[\epsilon \int_0^T Z_{\mathrm{N}}(t)\beta_{\mathrm{N}}(t)dW_{\mathrm{N}}\right] + O\left(\epsilon^2\right) \\
&= \mathbf{E}\left[\left(\epsilon \int_0^T Z_{\mathrm{V}}(t)\beta_{\mathrm{V}}(t)dW_{\mathrm{V}}\right)^2\right] \\
&+ \mathbf{E}\left[\left(\epsilon \int_0^T Z_{\mathrm{M}}(t)\beta_{\mathrm{M}}(t)dW_{\mathrm{M}}\right)^2\right] \\
&+ \mathbf{E}\left[\left(\epsilon \int_0^T Z_{\mathrm{N}}(t)\beta_{\mathrm{N}}(t)dW_{\mathrm{N}}\right)^2\right] + O\left(\epsilon^2\right),
\end{aligned} \tag{6.74}$$

where by the Itō isometry,

$$\begin{aligned}
&= \epsilon^2 \int_0^T Z_{\mathrm{V}}^2(t)\beta_{\mathrm{V}}^2(t)dt \\
&+ \epsilon^2 \int_0^T Z_{\mathrm{M}}^2(t)\beta_{\mathrm{M}}^2(t)dt \\
&+ \epsilon^2 \int_0^T Z_{\mathrm{N}}^2(t)\beta_{\mathrm{N}}^2(t)dt \; + O\left(\epsilon^2\right).
\end{aligned} \tag{6.75}$$

The ISI variance of a neural oscillator subject to Gaussian voltage perturbations was first derived by [17], who took a distinct approach which relied on higher order expansions of θ. The more general approach taken in this project facilitates an extension of [17] to more complex perturbations. In particular, by recognizing that SDE models of neural oscillators with weak noise (like the Morris-Lecar model defined in section 5.1 for channel numbers exceeding 30,000) can be expressed as a series of stochastic perturbations acting on a deterministic model, the explicit calculation of ISI distributions for more detailed weak-noise SDE models is made possible. In the next section stochastic Morris-Lecar perturbations are defined and the associated ISI

statistics are subsequently calculated.

6.4 Weak Stochastic Morris-Lecar Perturbations of Neural Oscillators

It was established in section 5.1 that for sufficiently large channel numbers the stochastic Morris Lecar model may be expressed in the form of an SDE,

$$\begin{bmatrix} dV \\ dX \\ dY \end{bmatrix} = \begin{bmatrix} F_V(V,X,Y) \\ F_X(V,X,Y) \\ F_Y(V,X,Y) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \epsilon_X \sigma_X(V,X,Y) dW_X(t) \\ \epsilon_Y \sigma_Y(V,X,Y) dW_Y(t) \end{bmatrix}, \tag{6.76}$$

whose solution (when ϵ_X and ϵ_Y are zero) is a stable, T-periodic limit cycle $\gamma(\theta(t))$. Instead of interpreting this as a SDE, we can think of it as a deterministic system perturbed by scaled Gaussian white noise where,

$$\vec{P}^*(V,X,Y,t) = \begin{bmatrix} 0 \\ \epsilon_X \sigma_X(V,X,Y)\xi_X(t) \\ \epsilon_Y \sigma_Y(V,X,Y)\xi_Y(t) \end{bmatrix}. \tag{6.77}$$

If,

$$\epsilon = \max\left(\epsilon_X, \epsilon_Y\right), \tag{6.78}$$

then for sufficiently small ϵ the system will stay near the limit cycle solution $\gamma(\theta(t))$,

$$\vec{P}^*(t) = \begin{bmatrix} 0 \\ \epsilon_X \sigma_X(\gamma(\theta(t)))\xi_X(t) \\ \epsilon_Y \sigma_Y(\gamma(\theta(t)))\xi_Y(t) \end{bmatrix}. \tag{6.79}$$

The pre-factor of ϵ once again enables the substitution of the zeroth order approximation of θ, and gives the weak stochastic Morris-Lecar perturbations,

$$\vec{P}^*(t) = \begin{bmatrix} 0 \\ \epsilon_X \sigma_X(\gamma(t))\xi_X(t) \\ \epsilon_Y \sigma_Y(\gamma(t))\xi_Y(t) \end{bmatrix}. \tag{6.80}$$

6.4.1 Interspike Interval Distribution of Neural Oscillator Subject to Stochastic Morris-Lecar Perturbations

If $\vec{Z}$ is obtained from the adjoint equation (eq. 6.19), with F representing the deterministic Morris-Lecar model, and the weak stochastic Morris-Lecar perturbations defined in equation 6.80 are substituted both into equation 6.67 we get,

$$R\left[\vec{P}^*(t)\right] = \int_0^T \vec{Z}(t) \cdot \vec{P}^*(t)dt + O\left(\epsilon^2\right). \tag{6.81}$$

or,

$$\begin{aligned} \hat{T}\left[\vec{P}^*(t)\right] &= T - \int_0^T \vec{Z}(t) \cdot \vec{P}^*(t)ds + O\left(\epsilon^2\right) \\ &= T - \epsilon_X \int_0^T Z_X(t)\, \sigma_X(\gamma(t))dW_X \\ &\quad - \epsilon_Y \int_0^T Z_Y(t)\, \sigma_Y(\gamma(t))dW_Y + O\left(\epsilon^2\right), \end{aligned} \tag{6.82}$$

which gives the interspike interval distribution of the weak noise Morris-Lecar Neuron to first order. Repeating the same calculations done in equations 6.72 and 6.73 it is then straightforward to find the mean,

$$\mathbf{E}\left[\hat{T}\left[\vec{P}^*(t)\right]\right] = T, \tag{6.83}$$

and variance,

$$\begin{aligned}\mathrm{Var}\left[\hat{T}\left[\vec{P}^*(t)\right]\right] &= \mathbf{E}\left[\left(\epsilon_X \int_0^T Z_X(t)\ \sigma_X(\gamma(t))dW_X\right)^2\right] \\ &+ \mathbf{E}\left[\left(\epsilon_Y \int_0^T Z_Y(t)\ \sigma_Y(\gamma(t))dW_Y\right)^2\right] + O\left(\epsilon^2\right) \\ &= \epsilon_X^2 \int_0^T Z_X^2(t)\ \sigma_X^2(\gamma(t))dt \\ &+ \epsilon_Y^2 \int_0^T Z_Y^2(t)\ \sigma_Y^2(\gamma(t))dt + O\left(\epsilon^2\right),\end{aligned} \tag{6.84}$$

where the independence of the noise sources as well as Itō's isometry were utilized. Since each σ was derived in section 5.1, the limit cycle of the deterministic system, γ, can be found numerically, and $\vec{Z}$ can be found by solving equation 6.19, the ISI variance plane described by this equation can be readily calculated and visualized.

To see the relationship between the expression above and other sections, note that the integrals are constants and that each ϵ can be expressed in terms of channel number,

$$\begin{aligned}\mathrm{Var}\left[ISI_{N_\mathrm{M},N_\mathrm{N}}\right] &= \frac{1}{N_\mathrm{M}} \int_0^T Z_X^2(t)\ \sigma_X^2(\gamma(t))dt \\ &+ \frac{1}{N_\mathrm{N}} \int_0^T Z_Y^2(t)\ \sigma_Y^2(\gamma(t))dt + O\left(\epsilon^2\right),\end{aligned} \tag{6.85}$$

which gives the ISI variance of a given cell model in terms of its channel numbers. Note that, as discussed in previous sections, the variance is to second order and that there are unspecified terms of order 2. This is likely the result of the discrepancy shown in Figure 6.1. In the next section this line of reasoning is carried out, and the ISI variance plane is compared to the SDE model in the small noise region of channel number space.

6.4.2 Numerical Convergence of SDE Model to Weak Perturbation Plane

Numerically calculating the integrals in equation 6.84 requires three steps. The first is to find the deterministic limit cycle as a function of t, $\gamma(t)$. To obtain the limit cycle, the deterministic channel state equations were propagated using Euler's method (so that they would be consistent with the SDE simulation which uses Euler-Maruyama) with a time step of 0.004 ms while the voltage equation was propagated using the exact solution. Once a long chain of spikes had been generated, one ISI was taken out from the group and used for the remainder of the calculation. Both σ_X and σ_Y were used in the same form that they were derived in section 5.1. Lastly, the iPRC was obtained through backwards integration (for stability) of the adjoint equation (derived in [16]) using the same time step 0.004 ms,

$$\frac{d}{dt}\vec{Z}(t) = -DF^T(\gamma(t))\vec{Z}(t), \tag{6.86}$$

where F is the vector form of the deterministic part of equation 6.76. Once each of these quantities were obtained, they were numerically integrated with the trapezoidal method over the time of the ISI which was 114.2 ms.

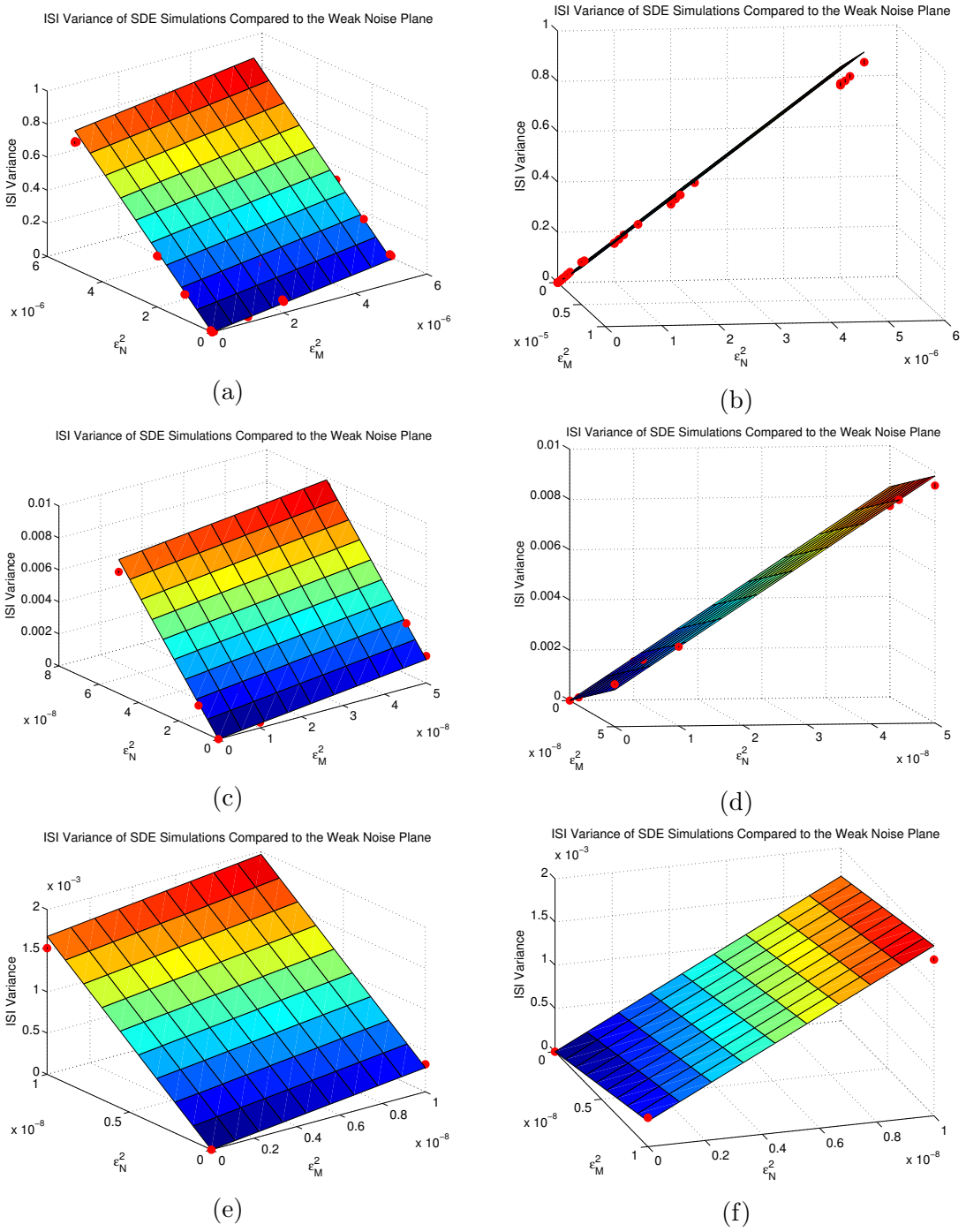


Figure 6.1: The x and y axis are related to the number of channels by $\epsilon^2 = 1/N$. Red points are the result of numerical simulations of the SDE model defined in section 5.1 using a time step of 0.004 ms. Dots located on $\epsilon = 0$ lines had the respective stochastic channel equation replaced by a deterministic one, representing an infinite channel limit. The origin represents a fully deterministic model. The plane is defined by equation 6.84, where $\gamma(t)$ is found through numerical simulation of the deterministic version of the SDE model using the same time step, both σ terms are used as defined in section 5.1, and $\vec{Z}$ is found by solving equation 6.86. The X integral in equation 6.86 was found to be 2.97×10^4, while the Y integral was found to be 1.31×10^5. Numerically fitting the points generated by the SDE model to a plane yields distinct coefficients with 95% confidence 3.358×10^4 (3.319×10^4, 3.396×10^4) and 1.535×10^5 (1.531^5, 1.539×10^5), which establishes that the two expressions of ISI variance are close but not equivalent. The source of the discrepancy is likely to be the unspecified second order terms in equation 6.85. The terms can be found by expanding to a higher order and performing the derivation again, but the result would contain derivatives of the iPRC, which are not readily obtainable.

Chapter 7

Statistical Convergence Between Models

Since this study relies heavily on the use of stochastic simulation to illuminate the Morris-Lecar neuron, it is essential that the uncertainty of the numerical simulations be quantified. Namely, some method must be derived which measures not only the sample mean, variance or CV of the ISI for a given system, but also the variability of each measurement. Additionally, since a central goal of the project is to numerically determine when progressive approximations of the exact model become valid in terms of these ISI statistics, a method must be developed to identify when the (uncertain) ISI statistics of two models converge in a statistically significant way. This chapter begins by finding computationally efficient approximations of the sample error, and then develops a method which uses the sample statistics to evaluate the convergence between two models.

Estimating the standard error of a sample variable using the brute force approach requires repeated calculation of the sample variable so that its variance can be assessed. Such a method is computationally expensive because if q ISIs are required and an accurate measure of the standard error requires w sample means, then the

standard error calculation requires qw ISIs. Given the high computational cost of each simulation, the brute force approach is not feasible, and some approximation must be developed.

The following chapter contains four parts. First, the standard error of the sample mean, variance, and CV are derived using the central limit theorem as well as a perturbative expansion. Next the theoretical equations which give the sample statistics will be re-expressed in terms of measurable, sample parameters, and then the estimation of the sample error is numerically validated. The chapter concludes by using the distribution of the sample statistics to construct a hypothesis test which uses ISI sample statistics to determine the significance of model convergence.

7.1 Estimating Error of Sample Mean, Variance, and CV

We can express the standard error of both the sample mean and variance through a simple application of the central limit theorem. These equations then define the sample CV of the ISI distribution which can be simplified through perturbative expansions. The expressions will be left in terms of unmeasurable parameters until they are re-expressed in the next section using sample statistics.

7.1.1 Notation and the Central Limit Theorem

Given an interspike interval random variable, X, we can take N i.i.d. samples denoted, $(X_1, ..., X_{\mathrm{N}}) \sim X$, where we know that $\mathbf{E}[X]$ and $\mathrm{Var}[X] < \infty$. If we define the sample mean of the ISI random variable as,

$$S_N = \frac{X_1 + ... + X_N}{N}, \tag{7.1}$$

then the central limit theorem gives convergence in distribution to a normal random variable,

$$\lim_{N\to\infty} S_N \xrightarrow{d} \mathcal{N}\left(\mathbf{E}[X], \frac{\mathrm{Var}[X]}{N}\right). \tag{7.2}$$

7.1.2 Error of Sample Mean in Terms of Variance

If we denote the expectation value of X as, $\mu = \mathbf{E}[X]$, then for sufficiently large N we can use the central limit theorem above to express the sample mean, $\bar{\mu} = S_{\mathrm{N}}$, in terms of the standard deviation, σ, and a normal random variable, Z, as,

$$\bar{\mu} = \mu + \frac{\sigma}{\sqrt{N}} Z + \xi. \tag{7.3}$$

The last term accounts for any error which results in the expression for finite N and is defined as,

$$\xi = \mu + \frac{\sigma}{\sqrt{N}} Z - \bar{\mu}, \tag{7.4}$$

where,

$$\xi = o\left(\frac{1}{\sqrt{N}}\right), \tag{7.5}$$

and

$$\mathbf{E}\left[\xi\right] = 0. \tag{7.6}$$

With these in mind we re-write equation 7.7 as,

$$\bar{\mu} = \mu + \frac{\sigma}{\sqrt{N}} Z + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.7}$$

To find the standard error of the sample mean we calculate the variance of $\bar{\mu}$,

$$\begin{aligned}
\mathrm{Var}[\bar{\mu}] &= \int (m - \mathbf{E}[\bar{\mu}])^2 f_{\bar{\mu}}(m) dm \\
&= \int (m - \mu)^2 f_{\bar{\mu}}(m) dm \\
&= \int \left(\mu + \frac{\sigma}{\sqrt{N}} m + o\left(\frac{1}{\sqrt{N}}\right) - \mu \right)^2 f_Z(m) dm \\
&= \int \left(\frac{\sigma}{\sqrt{N}} m + o\left(\frac{1}{\sqrt{N}}\right) \right)^2 f_Z(m) dm \\
&= \int \frac{\sigma^2}{N} m f_Z(m) + o\left(\frac{1}{N}\right) f_Z(m) dm \\
&= \frac{\sigma^2}{N} \int m f_Z(m) dm + o\left(\frac{1}{N}\right) \int f_Z(m) dm \\
&= \frac{\sigma^2}{N} \mathbf{E}[Z] + o\left(\frac{1}{N}\right) \\
&= \frac{\sigma^2}{N} + o\left(\frac{1}{N}\right).
\end{aligned} \tag{7.8}$$

We can now conclude that the standard error of the sample mean is,

$$\begin{aligned}
\mathrm{Std}[\bar{\mu}] &= \sqrt{\frac{\sigma^2}{N} + o\left(\frac{1}{N}\right)} \\
&= \frac{\sigma}{\sqrt{N}} + \frac{\sqrt{N}}{2\sigma} o\left(\frac{1}{N}\right) \\
&= \frac{\sigma}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).
\end{aligned} \tag{7.9}$$

7.1.3 Error of Sample Variance in Terms of Fourth Moment and Variance

To find the error in the sample variance, we use the ISI random variable, X, to define a new variable, $Y = (X - \mu)^2$. The new variable represents the squared deviations of X from its expectation value, and the mean of Y is equal to the variance of X. So, we can use the central limit theorem (which is expressed in terms of means) to see how the sample mean of Y converges to its mean, or, in other words, how the sample

variance of X converges to the variance of X. Proceeding as described, we define Y as,

$$Y = (X - \mu)^2. \tag{7.10}$$

If we refer to the expectation of Y by μ_Y, then from the definition above it should be clear that,

$$\mathbf{E}[Y] = \mu_Y = \mathrm{Var}[X]. \tag{7.11}$$

Note that I will add subscripts, μ_Y, for clarity when necessary, and that any variable without a subscript should be understood to be in terms of X. Proceeding with the central limit theorem, we define,

$$S'_N = \frac{Y_1 + ... + Y_N}{N}, \tag{7.12}$$

and apply the theorem to yield that,

$$\lim_{N\to\infty} S'_N \xrightarrow{d} \mathcal{N}\left(\mu_Y, \frac{\sigma_Y}{N}\right). \tag{7.13}$$

Now we may express the sample mean of Y with a new Gaussian random variable, Z', which is assumed to be approximately independent of Z (to simplify calculations in the near future),

$$\bar{\mu}_Y = \mu_Y + \frac{\sigma_Y}{\sqrt{N}} Z' + \xi_Y. \tag{7.14}$$

This equation takes the same form as equation 7.7, and ξ_Y is also defined by equations 7.4, 7.5, and 7.6. So, using the same logic we used to get equation 7.8 gives us the variance of Y's sample mean, and therefore the variance of X's sample variance,

$$\mathrm{Var}[\bar{\mu}_Y] = \mathrm{Var}\left[\bar{\sigma}^2_X\right] = \frac{\sigma^2_Y}{N} + o\left(\frac{1}{N}\right). \tag{7.15}$$

To illuminate the expression above we next express σ_Y^2 in terms of functions of X,

$$\begin{aligned}
\sigma_Y^2 &= \int (m-\mu_Y)^2 f_Y(m)dm \\
&= \int (m-\sigma_X^2)^2 f_Y(m)dm \\
&= \int ((m-\mu_X)^2-\sigma_X^2)^2 f_X(m)dm \\
&= \int \left[(m-\mu_X)^4 - 2\sigma_X^2(m-\mu_X)^2 + \sigma_X^4\right] f_X(m)dm \\
&= \int (m-\mu_X)^4 f_X(m)dm - 2\sigma_X^4 + \sigma_X^4 \\
&= \int (m-\mu_X)^4 f_X(m)dm - \sigma_X^4 \\
&= m_X^4 - \sigma_X^4,
\end{aligned} \tag{7.16}$$

where m_X^4 is the fourth central moment of X. As a result, we can re-write equation 7.15 in terms of X,

$$\text{Var}\left[\bar{\sigma}_X^2\right] = \frac{m_X^4 - \sigma_X^4}{N} + o\left(\frac{1}{N}\right), \tag{7.17}$$

or

$$\text{Std}\left[\bar{\sigma}_X^2\right] = \sqrt{\frac{m_X^4 - \sigma_X^4}{N}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.18}$$

Additionally, plugging the new result from 7.16 into equation 7.14 yields,

$$\begin{aligned}
\bar{\sigma}^2 &= \sigma^2 + \frac{\sqrt{m^4-\sigma^4}}{\sqrt{N}}Z' + \xi_Y \\
&= \sigma^2 + \sigma^2\frac{\sqrt{\frac{m^4}{\sigma^4}-1}}{\sqrt{N}}Z' + o\left(\frac{1}{\sqrt{N}}\right) \\
&= \sigma^2\left(1 + \frac{\sqrt{\gamma+2}}{\sqrt{N}}Z'\right) + o\left(\frac{1}{\sqrt{N}}\right),
\end{aligned} \tag{7.19}$$

where the kurtosis, γ, is defined as $\gamma = \frac{m^4}{\sigma^4} - 3$. From this expression we can find the

sample standard deviation,

$$\begin{aligned}\bar{\sigma} &= \sqrt{\sigma^2 + \sigma^2 \frac{\sqrt{\gamma+2}}{\sqrt{N}} Z' + o\left(\frac{1}{\sqrt{N}}\right)} \\ &= \sigma + \frac{1}{2\sigma}\left(\sigma^2 \frac{\sqrt{\gamma+2}}{\sqrt{N}} Z' + o\left(\frac{1}{\sqrt{N}}\right)\right) \\ &= \sigma + \sigma \frac{\sqrt{\gamma+2}}{2\sqrt{N}} Z' + o\left(\frac{1}{\sqrt{N}}\right) \\ &= \sigma\left(1 + \frac{\sqrt{\gamma+2}}{2\sqrt{N}} Z'\right) + o\left(\frac{1}{\sqrt{N}}\right).\end{aligned} \tag{7.20}$$

7.1.4 Error of Sample CV in Terms of Kurtosis and CV

We have already found expressions for both $\bar{\mu}$ (equation 7.7) and $\bar{\sigma}$ (equation 7.19). Instead of applying the central limit theorem, as we have up to this point, we substitute the expressions we already have into $\bar{\rho} = \bar{\sigma}/\bar{\mu}$, and simplify through the development of a perturbative expansion.

We begin with expressions of the constituent variables (equations 7.7 and 7.19 both of which are in terms of X),

$$\bar{\mu} = \mu + \frac{\sigma}{\sqrt{N}} Z + o\left(\frac{1}{\sqrt{N}}\right), \quad \bar{\sigma} = \sigma\left(1 + \frac{\sqrt{\gamma+2}}{2\sqrt{N}} Z'\right) + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.21}$$

We can now use the above expressions to write $\bar{\rho}$, where Z and Z' are assumed to be independent,

$$\begin{aligned}\bar{\rho} &= \frac{\bar{\sigma}}{\bar{\mu}} \\ &= \left(\sigma\left(1 + \frac{\sqrt{\gamma+2}}{2\sqrt{N}} Z'\right) + o\left(\frac{1}{\sqrt{N}}\right)\right)\left(\frac{1}{\mu + \frac{\sigma}{\sqrt{N}} Z + o\left(\frac{1}{\sqrt{N}}\right)}\right).\end{aligned} \tag{7.22}$$

For simplicity, we introduce the parameter,

$$\epsilon = \frac{1}{\sqrt{N}}, \tag{7.23}$$

so that,

$$\begin{aligned}\bar{\rho} &= \left(\sigma\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'\right)+o(\epsilon)\right)\left(\frac{1}{\mu+\sigma\epsilon Z+o(\epsilon)}\right)\\ &= \left(\rho\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'\right)+o(\epsilon)\right)\left(\frac{1}{1+\rho\epsilon Z+o(\epsilon)}\right)\\ &= \left(\rho\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'\right)+o(\epsilon)\right)(1-\rho\epsilon Z+o(\epsilon)).\end{aligned} \tag{7.24}$$

The (presumed) independence of Z and Z' allows us to combine these two Gaussian deviates into a new Z'' with a variance that is the sum of the variances associated with Z and Z',

$$\begin{aligned}\bar{\rho} &= \left(\rho\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'\right)+o(\epsilon)\right)(1-\rho\epsilon Z+o(\epsilon))\\ &= \rho\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'\right)-\epsilon Z\rho^2\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'\right)+o(\epsilon)\\ &= \rho\left(1+\frac{\sqrt{\gamma+2}}{2}\epsilon Z'-\epsilon Z\rho\right)+o(\epsilon)\\ &= \rho\left(1+\epsilon\left(\frac{\sqrt{\gamma+2}}{2}Z'-Z\rho\right)\right)+o(\epsilon)\\ &= \rho\left(1+\left(\sqrt{\frac{\gamma+2+4\rho^2}{4}}\right)\epsilon Z''\right)+o(\epsilon)\\ &= \rho\left(1+\frac{\sqrt{\gamma+2+4\rho^2}}{2}\epsilon Z''\right)+o(\epsilon).\end{aligned} \tag{7.25}$$

We now write this equation in "central limit theorem form" to emphasize its resemblance to equations 7.7 and 7.19,

$$\bar{\rho} = \rho + \rho\frac{\sqrt{\gamma+2+4\rho^2}}{2\sqrt{N}}Z'' + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.26}$$

Following similar logic to that used to obtain equation 7.8 it is now easy to see that,

$$\mathrm{Std}\left[\bar{\rho}\right] = \rho\frac{\sqrt{\gamma+2+4\rho^2}}{2\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.27}$$

To derive a more useful expression, in the next section we re-express equation 7.27 in

terms of sample statistics which can be readily obtained from numerical experiments.

7.2 Estimating Error of Sample Mean, Variance, and CV in Terms of Measurable Quantities

In the previous section we used the central limit theorem to express sample statistics (equations 7.7 and 7.19) and manipulated them to express their standard error (equations 7.9 and 7.18). These elementary expressions paired with perturbation theory enabled us to write $\bar{\rho}$ in "central limit theorem form" and calculate $\mathrm{Std}\left[\bar{\rho}\right]$. In this section we take the equations that have been developed thus far, and express them in terms of measurable quantities.

7.2.1 Error of Sample Mean In Terms of Sample Variance

In the previous section the central limit theorem was used to derive equation 7.9 which expresses the standard error of the sample mean,

$$\mathrm{Std}[\bar{\mu}] = \frac{\sigma}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.28}$$

This equation cannot be used in its current form to determine the error of the sample mean because it requires knowledge of the true, underlying standard deviation of the ISI random variable, σ. Since the true standard deviation is presently inaccessible, it must be replaced by an estimate. The central limit theorem has showed that the sample standard deviation is a Gaussian distance from the true standard deviation,

$$\bar{\sigma} = \sigma + \sigma\frac{\sqrt{\gamma+2}}{2\sqrt{N}}Z' + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.29}$$

With this in mind we assume that our best guess, $\tilde{\sigma}$, of the true standard deviation, σ, is a Gaussian distance from the sample standard deviation, $\bar{\sigma}$,

$$\tilde{\sigma} = \bar{\sigma} + \bar{\sigma}\frac{\sqrt{\gamma+2}}{2\sqrt{N}}Z' + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.30}$$

Substituting our best guest of the standard deviation, $\tilde{\sigma}$, for the true standard deviation, σ, into equation 7.28 gives,

$$\text{Std}[\bar{\mu}] = \frac{\bar{\sigma}}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.31}$$

7.2.2 Error of Sample Variance In Terms of Fourth Sample Moment and Sample Variance

We would like to express equation 7.18 in terms of measurable quantities. So, our task is to approximate the fourth moment of X, m^4, and the squared variance of X, σ^4, in terms of sample statistics. We can already get an approximation for σ^4 by using equation 7.30, so all that needs to be done is to raise $\tilde{\sigma}$ to the fourth power and remove higher order terms. Proceeding as described,

$$\begin{aligned}\tilde{\sigma}^4 &= \tilde{\sigma}^2\tilde{\sigma}^2 \\ &= \tilde{\sigma}^2\left(\bar{\sigma}^2 + \bar{\sigma}^2\frac{\sqrt{\gamma+2}}{\sqrt{N}}Z' + o\left(\frac{1}{\sqrt{N}}\right)\right) \\ &= \bar{\sigma}^4 + 2\bar{\sigma}^4\frac{\sqrt{\gamma+2}}{\sqrt{N}}Z' + o\left(\frac{1}{\sqrt{N}}\right).\end{aligned} \tag{7.32}$$

Since equation 7.17 is over N, we only need the zeroth order term to maintain first order accuracy for the entire expression, so we rewrite the above as,

$$\tilde{\sigma}^4 = \bar{\sigma}^4 + O\left(\frac{1}{\sqrt{N}}\right). \tag{7.33}$$

Now that an estimate of σ^4 has been developed, we can estimate the fourth central moment of X, m^4, in much the same way that we estimated σ earlier this section. To begin, we first define a new random variable that has an expectation value equal to the fourth central moment of X, and then use the central limit theorem to see how the sample mean of the new variable converges to its mean (which is the fourth moment of X). From this we find an estimation of m^4 analogous to $\tilde{\sigma}$. We begin by defining a new random variable,

$$M = (X - \mu)^4, \tag{7.34}$$

where X is an ISI random variable and $\mu = \mathbf{E}[X]$. Using the partial sum,

$$S''_{\mathrm{N}} = \frac{M_1 + ... + M_{\mathrm{N}}}{N}, \tag{7.35}$$

the central limit theorem gives,

$$\lim_{N\to\infty} S''_{\mathrm{N}} \xrightarrow{d} \mathcal{N}\left(\mathbf{E}[M], \frac{\mathrm{Var}[M]}{N}\right), \tag{7.36}$$

or

$$\bar{\mu}_{\mathrm{M}} = \mu_{\mathrm{M}} + \frac{\sigma_{\mathrm{M}}}{\sqrt{N}} Z + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.37}$$

Arguing the same way that we did to get equation 7.30 (and avoiding the lengthy and unnecessary calculation of σ_{M}) we can approximate $\mu_{\mathrm{M}} = m_X^4$ with a best guess variable, $\tilde{m}^4$,

$$\tilde{\mu_{\mathrm{M}}} = \bar{\mu_{\mathrm{M}}} + \frac{\sigma_{\mathrm{M}}}{\sqrt{N}} Z + o\left(\frac{1}{\sqrt{N}}\right), \tag{7.38}$$

or in terms of X,

$$\tilde{m}_X^4 = \bar{m}_X^4 + \frac{\sigma_{\mathrm{M}}}{\sqrt{N}} Z + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.39}$$

Since we eventually intend on plugging this expression into equation 7.17, which has

a pre-factor of N^{-1}, we only need zeroth order terms. So,

$$\tilde{m}_X^4 = \bar{m}_X^4 + O\left(N^{-1/2}\right). \tag{7.40}$$

Plugging equation 7.40 and 7.33 into equation 7.18 yields,

$$\text{Std}\left[\bar{\sigma}^2\right] = \sqrt{\frac{\bar{m}^4 - \bar{\sigma}^4}{N}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.41}$$

7.2.3 Error of Sample CV in Terms of Sample Kurtosis and CV

We will now express the empirical estimate for the standard deviation of the coefficient of variation (equation 7.27) in terms of measurable parameters. Arguing in the same ways as we did in equation 7.30, we can use equation 7.26 to express our best guess of the CV,

$$\tilde{\rho} = \bar{\rho} + \bar{\rho}\frac{\sqrt{\gamma + 2 + 4\rho^2}}{2\sqrt{N}}Z + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.42}$$

Similarly, we can see from equations 7.30 and 7.39 that γ can be approximated by,

$$\tilde{\gamma} = \bar{\gamma} + O(N^{-1/2}). \tag{7.43}$$

As in the last case, the pre-factor of $N^{-1/2}$ enables us to substitute the zeroth order term of the approximation and still maintain first order accuracy for the entire expression,

$$\tilde{\rho} = \bar{\rho} + \bar{\rho}\frac{\sqrt{\bar{\gamma} + 2 + 4\bar{\rho}^2}}{2\sqrt{N}}Z + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.44}$$

Finally, upon substituting $\tilde{\rho}$ in equation 7.44 in place of ρ in equation 7.27 we realize that the pre-factor of $N^{-1/2}$ in equation 7.27 cancels the higher order terms. So, we

are left with,

$$\mathrm{Std}\left[\bar{\rho}\right] = \bar{\rho}\frac{\sqrt{\bar{\gamma}+2+4\bar{\rho}^2}}{2\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right). \tag{7.45}$$

Equation 7.45 can be compared with its parent, equation 7.27, which expressed in terms of true parameters, ρ and γ.

7.3 Numerical Validation of Asymptotic Approximations

To validate the asymptotic approximations derived in the previous sections, we will show that they are within one standard deviation of the brute force method and that after some number of samples they scale in the same way. The first test shows that any given guess produced by the approximation will be close to the "real" underlying answer, while the second test shows that although we have truncated higher order terms, we nevertheless retain the proper result after a feasible number of samples.

7.3.1 Method of Validation

The brute force method and the asymptotic approximation will both be applied to the ISI distribution produced by the piecewise constant propensity algorithm with 1000 channels of each type. After about one million ISI samples were simulated with this configuration on a high performance computing cluster, a kernel smoothing function estimate (Matlab function ksdensity [2]) was applied to the ISI samples to yield an experimental probability density function (PDF). Using the PDF, the experimental cumulative density function (CDF) was found through numerical integration, and then the inverse transform method was applied to create an algorithm which samples from a smoothed, experimental, ISI distribution. With a general sampling method, hundreds of millions of artificial samples can be accessed and utilized for the numerical

validation of the standard error approximation. Since the experimental ISI samples which determined the sampler were so numerous, a distribution of real ISIs generated through the computing cluster is indecipherable by the eye from ISI samples generated through the artificial sampling method.

To perform the tests we will consider a matrix of ISI samples taking the following form,

$$I_k = \begin{bmatrix} ISI_{1,1} & \cdots & ISI_{1,10000} \\ \vdots & \cdots & \vdots \\ ISI_{k,1} & \cdots & ISI_{k,10000} \end{bmatrix}. \tag{7.46}$$

Statistical operators, S, like mean, variance, and CV act column wise on this matrix to yield a row vector,

$$S[I_k] = \begin{bmatrix} S\begin{bmatrix} ISI_{1,1} \\ \vdots \\ ISI_{k,1} \end{bmatrix} & \cdots & S\begin{bmatrix} ISI_{1,10000} \\ \vdots \\ ISI_{k,10000} \end{bmatrix} \end{bmatrix}. \tag{7.47}$$

The application of an additional statistical operator, S' on the row vector gives a number,

$$S'[S[I_k]] = S'\left[\begin{bmatrix} S\begin{bmatrix} ISI_{1,1} \\ \vdots \\ ISI_{k,1} \end{bmatrix} & \cdots & S\begin{bmatrix} ISI_{1,10000} \\ \vdots \\ ISI_{k,10000} \end{bmatrix} \end{bmatrix}\right]. \tag{7.48}$$

So, to find how the standard error of a sample statistic, S, diminishes as more samples, k, are taken using the brute force method, for each k we simply calculate, $\mathrm{Std}\,[S[I_k]]$. Note that while the number of samples which determine the sample statistic, S, grows with k, the number of samples which determines the standard error of S remains fixed, 10000. As k grows the brute force method (as well as its validation) requires $10000k$ samples, the number 10000 was chosen because it happens to be the largest number of columns with which the following tests could run on a personal computer.

Using the same matrix, I_k, we can apply the asymptotic approximation of the mean, variance, or CV (equation 7.31, 7.41, or 7.45) to each column to get a row vector of 10000 samples from the asymptotic approximation distribution (AAD). If we can establish that after some number of ISI samples, k_0, that the AAD is centered and tight around the brute force answer for all $k_0 < k$, then the replacement of the brute force method by the faster approximation will be justified. More specifically, we show numerically that after some k_0 the mean of the AAD converges to the brute force answer, and its standard deviation about that point is small.

7.3.2 Verification and Scope of Standard Error Approximation

By generating, I_k, many times across a range of k values using the numerical ISI sampler discussed above, both the brute force result, as well as the mean and variances of the AAD were calculated. The numerical experiment is depicted for mean, variance, and CV, in Figure 7.1. Beginning with the right column, we can see that in each case the mean of the AAD has converged to the brute force result after 20,000 samples, and that the AAD is tight around the mean. The left column shows that, in accordance with our approximations, the brute force method scales like $1/\sqrt{N}$. While the asymptotic approximation uses a $1/\sqrt{N}$ term, more complex behavior is observed when the number of ISI samples scales across small numbers. This is likely because the sample statistic terms within the pre factor of the $1/\sqrt{N}$ term are not very accurate, as they are derived from so few ISI samples (rows of I_k).

As a result of the data in Figure 7.1, we conclude that as long the asymptotic approximations are applied to at least 20,000 ISIs, error estimation of the mean, variance, and CV will be sufficiently close to a robust, brute force method. The development of these efficient error estimation techniques were essential for the numerical feasibility of large scale stochastic simulations. Using a brute force method in

the results section of this project would have required at least 100 times as many ISI samples, thereby making simulation, even the high performance computing cluster, unfeasible. While the approximation appears both useful and trustworthy, it is important to note that this validation has been done on the ISI distribution generated from the piecewise constant propensity approximation of the exact model with a specific channel configuration (M=1000, N=1000). A more rigorous study would remove any doubt that the validation is altered by changes in the channel configuration or simulation model (exact, SDE, etc.) by testing the approximations on many distinct ISI distributions. However, due to resource limitations, we make the assumption that our validation here is sufficient to prove that the approximation is true for all of the models and channel configurations used in the project.

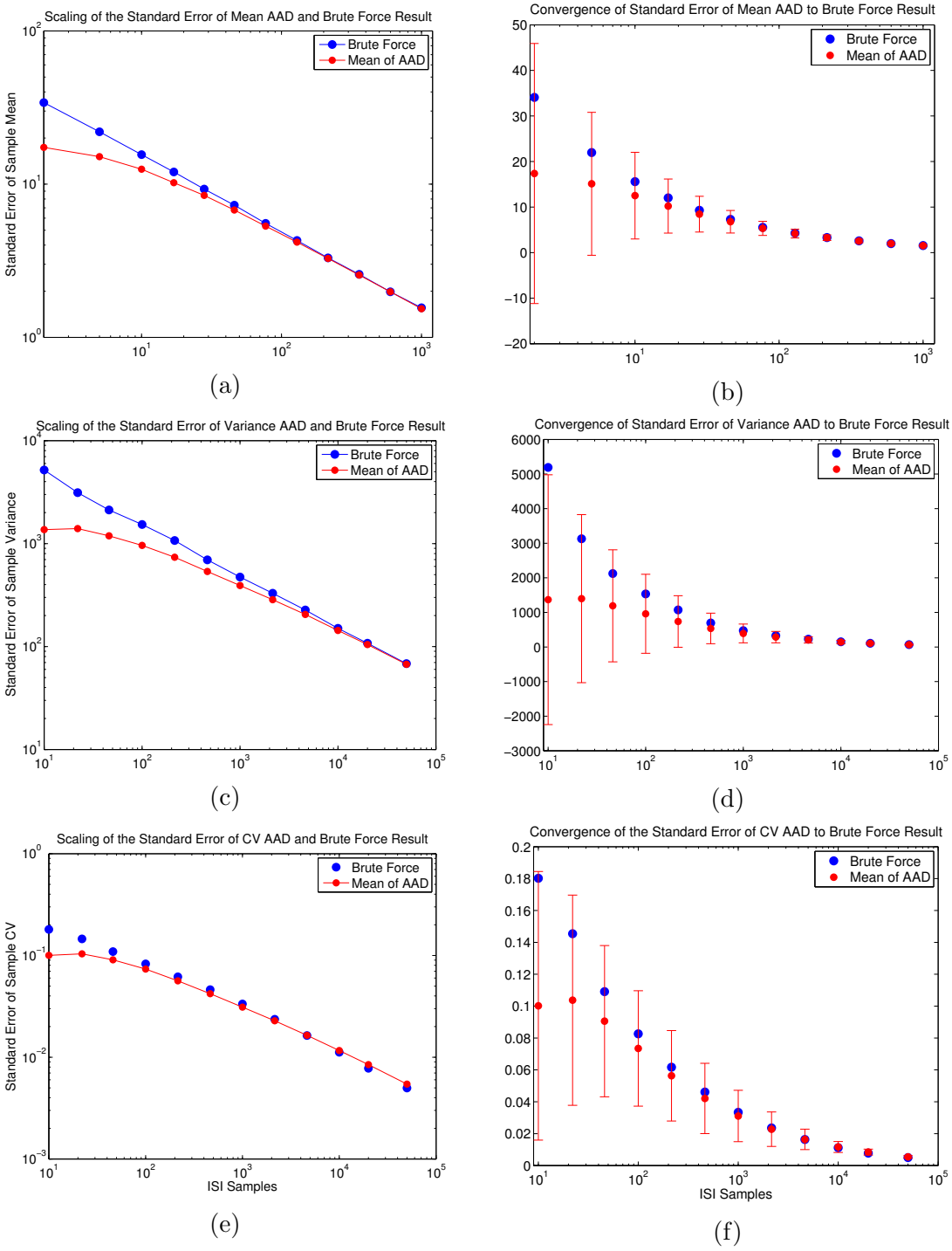


Figure 7.1: Standard error approximations of mean, variance, and CV outlined in section 7.2 are validated by applying a brute force and approximate method to large numbers of artificially generated ISIs. The artificial ISI sampler (described in section 7.3.1) was derived from 10^6 simulated ISIs using the PCPA model with 1000 M channels and 1000 N channels. Results indicate that the computationally efficient standard error approximation method reproduces the output of the more robust brute force method after 100 ISI samples for the sample mean, and 20,000 ISI samples for the standard error and CV (a,c,e) The mean of the asymptotic approximation distribution (AAD) of the ISI mean, variance, and CV scales, along with the brute force approximation, like $1/\sqrt{N}$, after about 1000 samples. (b,d,f) Re-expresses (a,c,e) on linear scale and establishes that all three measures of the AAD converge to the result of the brute force method after 20,000 samples (with the mean converging much earlier), indicating that a single draw from the AAD would replicate the brute force method with high probability.

7.4 Convergence of ISI Statistics

This project seeks to validate new approximations of the system defined in section 2.1.2 by determining when their ISI statistics converge to those of an exact model. Hence, a criteria for the convergence between the ISI statistics of two numerical implementations must be developed. Though it would be possible to extend these methods to the entire channel number plane, for the sake of simplicity, the methods in this section will only evaluate the convergence between two models along the $N_M = N_N$ line.

To consider the convergence between model A and model B, with matching channel configurations $N_N = N_M = C$, we denote the associated ISI distribution of each model as, ISI^{A_C} and ISI^{B_C}. After each model is simulated for a given amount of time, N_{A_C} spikes result from model A and N_{B_C} spikes result from model B. An expression for the sample statistics, which is developed in chapter 7 and numerically verified in Figure 7.1, shows that for a model Q, if $N_{Q_C} > 20000$ spikes result, then the following approximations hold with high accuracy,

$$\bar{\mu} \sim \mathcal{N}\left(\bar{\mu}, \frac{\bar{\sigma}^2}{N_{Q_C}}\right), \tag{7.49}$$

$$\bar{\sigma}^2 \sim \mathcal{N}\left(\bar{\sigma}^2, \frac{\bar{m}^4 - \bar{\sigma}^4}{N_{Q_C}}\right), \tag{7.50}$$

$$\bar{\rho} \sim \mathcal{N}\left(\bar{\rho}, \bar{\rho}\frac{\bar{\gamma} + 2 + 4\bar{\rho}^2}{4N_{Q_C}}\right). \tag{7.51}$$

where $\bar{\mu}, \bar{\sigma}^2, \bar{\rho}, \bar{\gamma}$, and $\bar{m}^4$ are the mean, variance, coefficient of variation (CV), kurtosis (defined as $\gamma = \frac{m^4}{\sigma^4} - 3$), and fourth moment of the sample distribution, $\text{ISI}^{Q_C}_{N_{Q_C}}$, respectively.

If we let $s_{N_{A_C}}$ denote a sample statistic from model A with channel configuration $N_N = N_M = C$ derived from N_{A_C} ISI samples, and similarly define $s_{N_{B_C}}$, then we can construct an illuminating hypothesis test. Under the null hypothesis that the two

underlying ISI distributions are equal,

$$H_0 : \mathrm{ISI}^{A_C} = \mathrm{ISI}^{B_C} \tag{7.52}$$

what is the probability of the observed (or more extreme) distance between the sample statistics?

Since the sample statistics are all well approximated by normal random variables with known parameters, we can calculate the associated p-value directly. Under the null hypothesis, the underlying distributions are equal, so we will take the parameters of the validated model, and substitute them in the unvalidated model. This section will outline the calculation for the sample mean where model A is the validated model. The case for the variance and CV are very similar. Proceeding with the case of the sample mean, if,

$$\bar{\mu}_{N_{A_C}} \sim \mathcal{N}\left(\bar{\mu}_{N_{A_C}}, \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{A_C}}\right), \quad \bar{\mu}_{N_{B_C}} \sim \mathcal{N}\left(\bar{\mu}_{N_{B_C}}, \frac{\bar{\sigma}^2_{N_{B_C}}}{N_{B_C}}\right), \tag{7.53}$$

then using the null hypothesis gives,

$$H_0 : \bar{\mu}_{N_{A_C}} \sim \mathcal{N}\left(\bar{\mu}_{N_{A_C}}, \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{A_C}}\right), \quad \bar{\mu}_{N_{B_C}} \sim \mathcal{N}\left(\bar{\mu}_{N_{A_C}}, \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{B_C}}\right). \tag{7.54}$$

Under the null hypothesis we have the distribution of the difference between the sample statistics of model A and model B,

$$H_0 : \quad \bar{\mu}_{N_{A_C}} - \bar{\mu}_{N_{B_C}} \sim \mathcal{N}\left(0, \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{A_C}} + \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{B_C}}\right), \tag{7.55}$$

which is standardized through a simple division,

$$H_0 : \quad \frac{\bar{\mu}_{N_{A_C}} - \bar{\mu}_{N_{B_C}}}{\sqrt{\frac{\bar{\sigma}^2_{N_{A_C}}}{N_{A_C}} + \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{B_C}}}} \sim \mathcal{N}(0, 1). \tag{7.56}$$

Now the p-value associated to model A, B, and channel configuration C is,

$$\mathrm{P}\left[-|X| > Z > |X| \;\middle|\; H_0 \right] = p_{N_{A_C}, N_{B_C}} \,, \tag{7.57}$$

where,

$$X = \frac{\bar{\mu}_{N_{A_C}} - \bar{\mu}_{N_{B_C}}}{\sqrt{\frac{\bar{\sigma}^2_{N_{A_C}}}{N_{A_C}} + \frac{\bar{\sigma}^2_{N_{A_C}}}{N_{B_C}}}}. \tag{7.58}$$

To evaluate the statistical significance of the convergence between model A and B across the $N_M = N_N = C$ line in the channel number plane, for each configuration C, the hypothesis test outlined above will be preformed, and the associated p-value will be calculated. Plotting the p-values as a function of channel number will illuminate how the statistical significance of the convergence changes as a function of channel numbers.

The figures 4.5 and 5.5 show the evaluation of two null hypotheses,

$$H_0 : \mathrm{ISI}^{\mathrm{Exact}_C} = \mathrm{ISI}^{\mathrm{PCPA}_C}, \tag{7.59}$$

and

$$H_0 : \mathrm{ISI}^{\mathrm{PCPA}_C} = \mathrm{ISI}^{\mathrm{SDE}_C}, \tag{7.60}$$

using same the calculation which derived expression 7.57 to test the convergence between the mean, variance, and CV of each model. Figure 4.5 shows p-values much smaller than the standard cutoff of 0.05. Therefore, we can confidently reject the null hypothesis in the first case, and conclude that the Exact model, as developed in this project, does not converge to the PCPA model. On the other hand, Figure 5.5 shows p-values which are consistently above the standard cutoff value, so we do not have sufficient evidence to reject the null hypothesis, and can conclude that the PCPA model converges to the SDE model in a statistically significant way when channel

numbers of each type exceed 400.

Bibliography

[1] David F Anderson, Bard Ermentrout, and Peter J Thomas. Stochastic representations of ion channel kinetics and exact stochastic simulation of neuronal dynamics. *Journal of computational neuroscience*, 38(1):67–82, 2015.

[2] Adrian W Bowman and Adelchi Azzalini. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations: The Kernel Approach with S-Plus Illustrations*. Oxford University Press, 1997.

[3] Ian C Bruce. Implementation issues in approximate methods for stochastic Hodgkin–Huxley models. *Annals of biomedical engineering*, 35(2):315–318, 2007.

[4] Ian C Bruce. Evaluation of stochastic differential equation approximation of ion channel gating models. *Annals of biomedical engineering*, 37(4):824–838, 2009.

[5] Lars Buesing, Johannes Bill, Bernhard Nessler, and Wolfgang Maass. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS computational biology*, 7(11):e1002211, 2011.

[6] Anthony N Burkitt. A review of the integrate-and-fire neuron model: I. homogeneous synaptic input. *Biological cybernetics*, 95(1):1–19, 2006.

[7] Anthony N Burkitt. A review of the integrate-and-fire neuron model: II. inhomogeneous synaptic input and network properties. *Biological cybernetics*, 95(2):97–112, 2006.

[8] Robert C Cannon, Cian O'Donnell, and Matthew F Nolan. Stochastic ion channel gating in dendritic neurons: morphology dependence and probabilistic synaptic activation of dendritic spikes. *PLoS computational biology*, 6(8):e1000886, 2010.

[9] Carson C Chow and John A White. Spontaneous action potentials due to channel fluctuations. *Biophysical Journal*, 71(6):3013–3021, 1996.

[10] John R Clay and Louis J DeFelice. Relationship between membrane excitability and single channel open-close kinetics. *Biophysical journal*, 42(2):151–157, 1983.

[11] James J Collins, Thomas T Imhoff, and Peter Grigg. Noise-enhanced tactile sensation. *Nature*, 1996.

[12] Kamran Diba, Henry A Lester, and Christof Koch. Intrinsic noise in cultured hippocampal neurons: experiment and modeling. *The Journal of neuroscience*, 24(43):9723–9733, 2004.

[13] Alan D Dorval and John A White. Channel noise is essential for perithreshold oscillations in entorhinal stellate neurons. *The Journal of neuroscience*, 25(43):10025–10028, 2005.

[14] John K Douglass, Lon Wilkens, Eleni Pantazelou, and Frank Moss. Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. *Nature*, 365(6444):337–340, 1993.

[15] Declan A Doyle, Joao Morais Cabral, Richard A Pfuetzner, Anling Kuo, Jacqueline M Gulbis, Steven L Cohen, Brian T Chait, and Roderick MacKinnon. The structure of the potassium channel: molecular basis of k+ conduction and selectivity. *science*, 280(5360):69–77, 1998.

[16] G Bard Ermentrout and David H Terman. *Mathematical foundations of neuroscience*, volume 64. Springer, 2010.

[17] GB Ermentrout, B Beverlin 2nd, T Troyer, and TI Netoff. The variance of phase-resetting curves. *Journal of computational neuroscience*, 31(2):185, 2011.

[18] Stewart N Ethier and Thomas G Kurtz. *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons, 2009.

[19] A Aldo Faisal and Simon B Laughlin. Stochastic simulations on the reliability of action potential propagation in thin axons. *PLoS computational biology*, 3(5):e79, 2007.

[20] A Aldo Faisal, Luc PJ Selen, and Daniel M Wolpert. Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303, 2008.

[21] Fernando R Fernandez, Jordan DT Engbers, and Ray W Turner. Firing dynamics of cerebellar purkinje cells. *Journal of neurophysiology*, 98(1):278–294, 2007.

[22] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445–466, 1961.

[23] Ronald F Fox. Stochastic versions of the Hodgkin-Huxley equations. *Biophysical journal*, 72(5):2068–2074, 1997.

[24] Ronald F Fox and Yan-nan Lu. Emergent collective behavior in large numbers of globally coupled independently stochastic ion channels. *Physical Review E*, 49(4):3421, 1994.

[25] Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 4. Springer Berlin, 1985.

[26] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[27] Joshua H Goldwyn, Nikita S Imennov, Michael Famulare, and Eric Shea-Brown. Stochastic differential equation models for ion channel noise in Hodgkin-Huxley neurons. *Physical Review E*, 83(4):041908, 2011.

[28] Joshua H Goldwyn and Eric Shea-Brown. The what and where of adding channel noise to the Hodgkin-Huxley equations. *PLoS computational biology*, 7(11):e1002247, 2011.

[29] Peter Hänggi. Stochastic resonance in biology how noise can enhance detection of weak signals and help improve biological information processing. *ChemPhysChem*, 3(3):285–290, 2002.

[30] Alan L Hodgkin and Andrew F Huxley. Action potentials recorded from inside a nerve fibre. *Nature*, 144(3651):710–711, 1939.

[31] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.

[32] Eugene M Izhikevich. *Dynamical systems in neuroscience*. MIT press, 2007.

[33] Staffan Johansson and Peter Arhem. Single-channel currents trigger action potentials in small cultured hippocampal neurons. *Proceedings of the National Academy of Sciences*, 91(5):1761–1765, 1994.

[34] Yoshiki Kuramoto. *Chemical oscillations, waves, and turbulence*. Courier Dover Publications, 2003.

[35] Thomas G Kurtz. Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2):344–356, 1971.

[36] Thomas G Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972.

[37] Harold Lecar and Ralph Nossal. Theory of threshold fluctuations in nerves: I. relationships between electrical noise and fluctuations in axon firing. *Biophysical journal*, 11(12):1048–1067, 1971.

[38] Sang-Gui Lee and Seunghwan Kim. Parameter dependence of stochastic resonance in the stochastic Hodgkin-Huxley neuron. *Physical Review E*, 60(1):826, 1999.

[39] Jacob E Levin and John P Miller. Broadband neural encoding in the cricket cereal sensory system enhanced by stochastic resonance. *Nature*, 380(6570):165–168, 1996.

[40] Daniele Linaro, Marco Storace, and Michele Giugliano. Accurate and fast simulation of channel noise in conductance-based model neurons by diffusion approximation. *PLoS computational biology*, 7(3):e1001102, 2011.

[41] Mark D McDonnell and Derek Abbott. What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology. *PLoS computational biology*, 5(5):e1000348, 2009.

[42] Hiroyuki Mino, Jay T Rubinstein, and John A White. Comparison of algorithms for the simulation of action potentials with stochastic sodium channels. *Annals of biomedical engineering*, 30(4):578–587, 2002.

[43] Catherine Morris and Harold Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical journal*, 35(1):193–213, 1981.

[44] Erwin Neher and Bert Sakmann. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature*, 260, 1976.

[45] Jay M Newby, Paul C Bressloff, and James P Keener. Breakdown of fast-slow analysis in an excitable system with channel noise. *Physical review letters*, 111(12):128101, 2013.

[46] Saak V Ovsepian and David D Friel. The leaner p/q-type calcium channel mutation renders cerebellar purkinje neurons hyper-excitable and eliminates ca2+-na+ spike bursts. *European Journal of Neuroscience*, 27(1):93–103, 2008.

[47] Khashayar Pakdaman, Michele Thieullen, Gilles Wainrib, et al. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Advances in Applied Probability*, 42(3):761–794, 2010.

[48] Dale Purves, GJ Augustine, D Fitzpatrick, WC Hall, AS LaMantia, JO McNamara, and LE White. Neuroscience. *De Boeck, Sinauer, Sunderland, Mass*, 2008.

[49] Hugh PC Robinson and Nobufumi Kawai. Injection of digitally synthesized synaptic conductance transients to measure the integrative properties of neurons. *Journal of neuroscience methods*, 49(3):157–165, 1993.

[50] JT Rubinstein. Threshold fluctuations in an n sodium channel model of the node of ranvier. *Biophysical journal*, 68(3):779–785, 1995.

[51] Antti Saarinen, Marja-Leena Linne, and Olli Yli-Harja. Stochastic differential equation model for cerebellar granule cell excitability. *PLoS computational biology*, 4(2):e1000004, 2008.

[52] Gerhard Schmid, Igor Goychuk, and P Hänggi. Stochastic resonance as a collective property of ion channel assemblies. *EPL (Europhysics Letters)*, 56(1):22, 2001.

[53] Gerhard Schmid, Igor Goychuk, and Peter Hänggi. Membrane clusters of ion channels: size effects for stochastic resonance. In *Statistical Mechanics of Complex Networks*, pages 195–206. Springer, 2003.

[54] Elad Schneidman, Barry Freedman, and Idan Segev. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural computation*, 10(7):1679–1703, 1998.

[55] Christof J Schwiening. A brief historical perspective: Hodgkin and Huxley. *The Journal of physiology*, 590(11):2571–2575, 2012.

[56] Andrew A Sharp, Michael B O'Neil, LF Abbott, and Eve Marder. Dynamic clamp: computer-generated conductances in real neurons. *Journal of neurophysiology*, 69:992–992, 1993.

[57] Erik Skaugen and Lars Walløe. Firing behaviour in a stochastic nerve membrane model based upon the Hodgkin—Huxley equations. *Acta Physiologica Scandinavica*, 107(4):343–363, 1979.

[58] Peter N Steinmetz, Amit Manwani, Christof Koch, Michael London, and Idan Segev. Subthreshold voltage noise due to channel fluctuations in active neuronal membranes. *Journal of computational neuroscience*, 9(2):133–148, 2000.

[59] Adam F Strassberg and Louis J DeFelice. Limitations of the Hodgkin-Huxley formalism: Effects of single channel kinetics on transmembrane voltage dynamics. *Neural Computation*, 5(6):843–855, 1993.

[60] Rafael D Vilela and Benjamin Lindner. Comparative study of different integrate-and-fire neurons: spontaneous activity, dynamical response, and stimulus-induced correlation. *Physical Review E*, 80(3):031909, 2009.

[61] John Von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. *Automata studies*, 34:43–98, 1956.

[62] John Von Neumann. *The computer and the brain*. Yale University Press, 2012.

[63] Joy T Walter, Karina Alvina, Mary D Womack, Carolyn Chevez, and Kamran Khodakhah. Decreases in the precision of purkinje cell pacemaking cause cerebellar dysfunction and ataxia. *Nature neuroscience*, 9(3):389–397, 2006.

[64] Maosheng Wang, Zhonghuai Hou, and Houwen Xin. Double-system-size resonance for spiking activity of coupled Hodgkin–Huxley neurons. *ChemPhysChem*, 5(10):1602–1605, 2004.

[65] John A White, Ruby Klink, Angel Alonso, and Alan R Kay. Noise from voltage-gated ion channels may influence neuronal dynamics in the entorhinal cortex. *Journal of neurophysiology*, 80(1):262–269, 1998.

[66] John A White, Jay T Rubinstein, and Alan R Kay. Channel noise in neurons. *Trends in neurosciences*, 23(3):131–137, 2000.

[67] John A White, Jay T Rubinstein, and Hiroyuki Mino. Response: Implementation issues in approximate methods for stochastic Hodgkin-Huxley models. *Annals of Biomedical Engineering*, 35(2):319–319, 2007.

[68] Shmuel Winograd and Jack D Cowan. *Reliable computation in the presence of noise*. MIT Press Cambridge, Mass., 1963.

[69] Yuguo Yu, Wei Wang, Jiafu Wang, and Feng Liu. Resonance-enhanced signal detection and transduction in the Hodgkin-Huxley neuronal systems. *Physical Review E*, 63(2):021907, 2001.