

NOISE DECOMPOSITION FOR STOCHASTIC HODGKIN-HUXLEY MODELS

by

SHUSEN PU

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Dissertation Advisor: Dr. Peter J. Thomas

Department of Mathematics, Applied Mathematics and Statistics

CASE WESTERN RESERVE UNIVERSITY

January, 2021

CASE WESTERN RESERVE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

We hereby approve the thesis of

Shusen Pu

Candidate for the Doctor of Philosophy degree*.

Prof. Peter J. Thomas

Committee Chair, Advisor

Prof. Daniela Calvetti

Professor of Mathematics

Prof. Erkki Somersalo

Professor of Mathematics

Prof. David Friel

External Faculty, Associate Professor of Neurosciences

Date: October 2, 2020

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

Table of Contents

Table of Contents	ii
List of Tables	v
List of Figures	vi
Acknowledgement	viii
Abstract	x
I Introduction and Motivation	1
1 Physiology Background	2
1.1 Single Cell Neurophysiology	2
1.2 Ion Channels	4
2 Foundations of the Hodgkin-Huxley Model	6
2.1 Membrane Capacitance and Reversal Potentials	7
2.2 The Membrane Current	8
2.3 The Hodgkin-Huxley Model	9
3 Channel Noise	14
3.1 Modeling Channel Noise	16
3.2 Motivation 1: The Need for Efficient Models	19
4 The Connection Between Variance of ISIs with the Random Gating of Ion Channels	22
4.1 Variability In Action Potentials	22
4.2 Motivation 2: Understanding the Molecular Origins of Spike Time Variability	23
II Mathematical Framework	31
5 The Deterministic 4-D and 14-D HH Models	32
5.1 The 4D Hodgkin-Huxley Model	35
5.2 The Deterministic 14D Hodgkin-Huxley Model	36

5.3	Relation Between the 14D and 4D Deterministic HH Models	37
5.4	Local Convergence Rate	47
6	Stochastic 14D Hodgkin-Huxley Models	51
6.1	Two Stochastic Representations for Stochastic Ion Channel Kinetics	52
6.2	Exact Stochastic Simulation of HH Kinetics	58
6.3	Previous Langevin Models	61
6.4	The 14×28 D HH Model	66
6.5	Stochastic Shielding for the 14D HH Model	71
III	Model Comparison	78
7	Pathwise Equivalence for a Class of Langevin Models	79
7.1	When are Two Langevin Equations Equivalent?	80
7.2	Map Channel-based Langevin Models to Fox and Lu's Model	84
8	Model Comparison	90
8.1	L_1 Norm Difference of ISIs	91
8.2	Two-sample Kolmogorov-Smirnov Test	96
IV	Applications of the 14×28D Langevin Model	102
9	Definitions, Notations and Terminology	103
9.1	The HH Domain	103
9.2	Interspike Intervals and First Passage Times	106
9.3	Asymptotic phase and infinitesimal phase response curve	114
9.4	Small-noise expansions	116
9.5	Iso-phase Sections	118
10	Noise Decomposition of the 14-D Stochastic HH Model	122
10.1	Assumptions for Decomposition of the Full Noise Model	124
10.2	Noise Decomposition Theorem	127
10.3	Proof of Theorem 4	131
11	Decomposition of the Variance of Interspike Intervals	139
11.1	Observations on σ_{ISI}^2 , and σ_{IPI}^2	142
11.2	Numerical Performance of the Decomposition Theorem	147

V	Conclusions and Discussion	156
12	Conclusions	157
12.1	Summary	157
13	Discussions	160
13.1	Discrete Gillespie Markov Chain Algorithms	160
13.2	Langevin Models	161
13.3	Model Comparisons	163
13.4	Stochastic Shielding Method	164
13.5	Which Model to Use?	165
13.6	Variance of Interspike Intervals	167
14	Limitations	173
A	Table of Common Symbols and Notations	177
B	Parameters and Transition Matrices	180
C	Diffusion Matrix of the 14D Model	183
	Bibliography	187

List of Tables

3.1	Number of Na ⁺ and K ⁺ channels	15
5.1	Map from 14D to 4D HH model	38
5.2	Map from 4D to 14D HH model	39
8.1	Model comparison	94
A.1	Common symbols and notations in this thesis (part I).	177
A.2	Common symbols and notations in this thesis (part II).	178
A.3	Common symbols and notations in this thesis (part III).	179
B.1	Parameters used for simulations in this thesis.	180

List of Figures

1.1	Diagrams of neurons	3
1.2	Schematic plot of a typical action potential	5
2.1	HH components	10
2.2	Sample trace of the deterministic HH model	13
3.1	Hodgkin-Huxley kinetics	17
4.1	Voltage recordings from intact Purkinje cells	24
4.2	Interdependent variables under current clamp	26
5.1	4D and 14D HH models	34
5.2	Convergence to the 4D submanifold	50
6.1	Stochastic shielding under voltage clamp	72
6.2	Stochastic shielding under current clamp	76
6.3	Robustness of edge-importance ordering under current clamp	77
7.1	Pathwise equivalency of 14D HH model and Fox and Lu’s model	89
8.1	Subunit model	95
8.2	Relative computational efficiency	97
8.3	Logarithm of the ratio of Kolmogorov-Smirnov test statistic	100
9.1	Example voltage trace for the 14-D stochastic HH model	107
9.2	Sample trace of the 14×28 D model	109
9.3	Example trajectory and the partition of threshold crossing	110
9.4	Sample trace on the limit cycle and its phase function	114
10.1	Approximate decomposition of interspike interval variance	124
11.1	Histogram of voltage recordings and number of spikes as a function of voltage threshold	144
11.2	Standard deviation of the interspike intervals and iso-phase intervals	146

11.3	Comparison of variance of ISIs (σ_{ISI}^2) and IPIs (σ_{IPI}^2)	148
11.4	Numerical performance of the decomposition for the ISI variance for Na ⁺ and K ⁺ kinetics	152
11.5	Overall performance of the decomposition theorem on ISI variance	154
13.1	Variance of ISI using different threshold conditions	169
13.2	Variance of ISI using different thresholds for large noise	171

Acknowledgement

This thesis represents not only my development in applied mathematics, it is also a milestone in six years of study at Case Western Reserve University (CWRU). My experience at CWRU has been a wonderful journey full of amazement. This thesis is the result of many experiences I have encountered at CWRU from dozens of remarkable individuals who I sincerely wish to acknowledge.

First and foremost, I would like to thank my advisor, Professor Peter J. Thomas for his support in both my academic career and personal life. He has been supportive even before I officially registered at CWRU when we met on the 2014 SIAM conference on the life sciences. I still remember him sketching the neural spike train and explaining to me the mathematical models for action potentials, which turned out to be the main focus of my dissertation. Ever since, Professor Thomas has supported me not only by teaching me knowledge in computational neuroscience over almost six years, but also academically and emotionally through the rough road to finish this thesis. Thanks to his great effort in walking me through every detail of the field and proofreading line by line of all my papers and thesis. And during the most difficult times when my father left me, he gave me the moral support and the time that I needed to recover. There are so many things fresh in my mind and I cannot express how grateful I am for his supervision and support.

I would also like to thank every professor who shared their knowledge and experience with me, Dr. Daniela Calvetti, Dr. Erkki Somersalo, Dr. Weihong Guo, Dr. Alethea Barbaro, Dr. Wanda Strychalski, Dr. Wojbor Woyczynski, and Dr. Longhua Zhao.

Especially, I want to thank my committee members Dr. Daniela Calvetti, Dr. Erkki Somersalo in the Department of Mathematics, Applied Mathematics and Statistics at CWRU, and Dr. David Friel from the Department of Neurosciences at the Case School of Medicine. The book on mathematical modeling written by Dr. Daniela Calvetti and Dr. Erkki Somersalo is my first formal encounter with mathematical models for dynamic systems and stochastic processes. Moreover, I have taken a series of computational oriented courses offered by Dr. Daniela Calvetti, where she taught me many programming skills and how to elegantly present my work. I would like to thank Dr. David Friel for his inspiring discussions during the past few years. All data recordings in this thesis come from Dr. David Friel's laboratory, and he taught me a lot of the biophysical background of this dissertation.

Moreover, I want to thank the department and National Science Foundation (grant DMS-1413770) for their support and sponsorship. The devoted faculty and staff in the department have been really helpful during my graduate study. I also want to thank my fellow graduate students for their friendship and support. Especially, Sumanth Nakkireddy, Alexander Strang, Isaac Oduro, Ben Li, Kai Yin, Alberto Bocchinfuso and Yue Zhang, who made this journey a lot of fun.

Finally, I would like to dedicate this thesis to my wife Yuanqi Xie and my parents. I feel extremely fortunate to have such a wonderful family. I could not have accomplished this without their love and support.

Stochastic Hodgkin-Huxley Models and Noise Decomposition

Abstract

by

SHUSEN PU

In this thesis, we present a natural 14-dimensional Langevin model for the Hodgkin-Huxley (HH) conductance-based neuron model in which each directed edge in the ion channel state transition graph acts as an independent noise source, leading to a 14×28 noise coefficient matrix. We show that (i) the corresponding 14D mean-field ordinary differential equation system is consistent with the classical 4D representation of the HH system; (ii) the 14D representation leads to a noise coefficient matrix that can be obtained cheaply on each timestep, without requiring a matrix decomposition; (iii) sample trajectories of the 14D representation are pathwise equivalent to trajectories of several existing Langevin models, including one proposed by Fox and Lu in 1994; (iv) our 14D representation give the most accurate interspike-interval distribution, not only with respect to moments but under both the L_1 and L_∞ metric-space norms; and (v) the 14D representation gives an approximation to exact Markov chain simulations that are as fast and as efficient as all equivalent models.

We combine the stochastic shielding (SS) approximation, introduced by Schmandt and Galán in 2012, with Langevin versions of the HH model to derive an analytic decomposition of the variance of the interspike intervals (ISI), based on the mean–return-time oscillator phase. We prove in theory, and demonstrate numerically, that in the limit of

small noise, the variance of the ISI decomposes linearly into a sum of contributions from each directed edge. Unlike prior analyses, our results apply to current clamp rather than voltage clamp conditions. Under current clamp, a stochastic conductance-based model is an example of a piecewise-deterministic Markov process. Our theory is exact in the limit of small channel noise. Through numerical simulations we demonstrate its applicability over a range from small to moderate noise levels. We show numerically that the SS approximation has a high degree of accuracy even for larger, physiologically relevant noise levels.

Part I

Introduction and Motivation

Chapter 1

Physiology Background

There is no scientific study more vital to man than the study of his own brain.

Our entire view of the universe depends on it.

– Francis Crick

1.1 Single Cell Neurophysiology

The human brain is the command center for the human nervous system, which controls most of the activities of the body, receiving and analyzing information from the body's sensory organs, and sending out decision information to the rest of the body. The human brain contains billions of nerve cells interconnected by trillions of synapses, that communicate with one another and with peripheral systems (sensory organs, muscles) through transient spikes in transmembrane voltages called “action potentials”. Fig. 1.1 shows three types of neurons. In a neuron, the dendrite receives inputs from other neurons, and the axon carries the neuronal output to other cells [22].

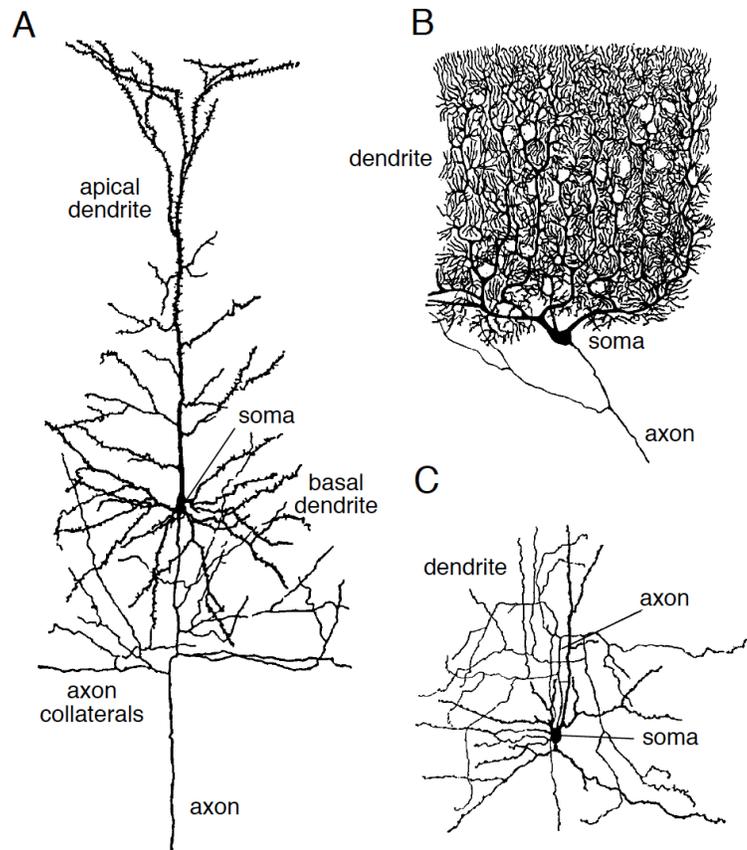


Figure 1.1: Morphology of neurons. A: a cortical pyramidal cell. B: a Purkinje cell of the cerebellum. C: a stellate cell of the cerebral cortex. Reproduced from [22], original drawing from [133]; figure from [27].

1.2 Ion Channels

There are a wide variety of pore-forming membrane proteins, namely ion channels, that allow ions, such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), and chloride (Cl^-), to pass through the cell. Ion channels control the flow of ions across the cell membrane by opening and closing in response to voltage changes and to both internal and external signals. Neurons maintain a voltage difference between the exterior and interior of the cell, which is called the membrane potential. Under resting conditions, a typical voltage across an neuron cell membrane is about -70 mV. Pumps spanning the cell membrane maintain a concentration difference that support this membrane potential. More specifically, under the resting state, the concentration of Na^+ is much higher outside a neuron than inside, while the concentration of K^+ is significantly higher inside the neuron than its extracellular environment [22].

An action potential occurs when the membrane potential at a specific location of the cell rapidly changes. Action potentials are generated by voltage-gated ion channels in the cell's membrane. These ion channels are shut when the membrane potential is near the resting potential and they rapidly open when the membrane potential increases to a threshold voltage, which leads a depolarization of the membrane potential [8]. During the depolarization, the sodium channels open, which produces a further rise in the membrane potential. The influx of sodium ions causes the polarity of the membrane to reverse, which rapidly leads to the inactivation of sodium channels and activation of potassium channels. There is a transient negative shift after an action potential has occurred, which is called the after hyperpolarization (AHP). Action potentials (or “spikes”) are generally thought to carry information via their timing (as opposed to their magnitude or duration). Figure 1.2 gives an schematic plot of a typical action potential.

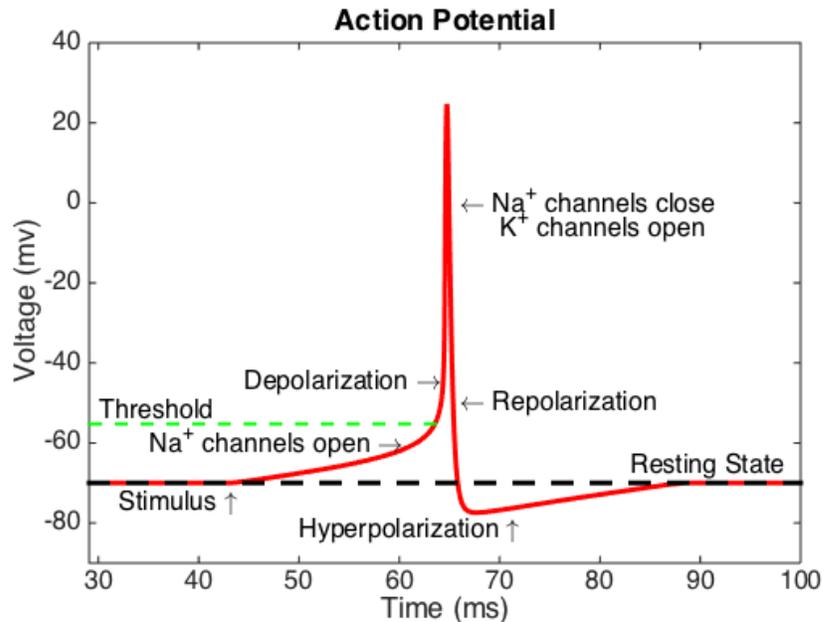


Figure 1.2: Schematic plot of a typical action potential on a cell membrane. The resting potential is ≈ -70 mV, and a stimulus is applied at time = 42 ms, which raises the membrane potential above a threshold voltage (≈ -55 mV in this case). After reaching the threshold, the membrane potential rapidly rises to a peak potential around 25 mV at time = 62 ms, which is followed by a rapid fall that overshoots -70 mV to a hyperpolarization state. After a resetting process, the membrane potential finally sets back to ≈ -70 mV around time = 88 ms. The spike itself is an all-or-none phenomenon, so information is coded not in the amplitude of a spike but in the timing of spikes.

The timing of action potentials is determined by a combination of deterministic factors external to a nerve cell (the transmembrane current it receives from its synapses, for example) and intrinsic stochastic factors such as the random gating of discrete populations of ion channels. Under experimental conditions in which a single nerve cell's external inputs are blocked, save for a constant applied driving current, the variability of action potential timing due to fluctuations in the states of several ion channels (sodium, potassium, and calcium channels being among the most prominent types) can be studied in the laboratory.

Chapter 2

Foundations of the Hodgkin-Huxley

Model

The Nobel Prize in Physiology or Medicine 1963 is awarded jointly to Sir John Carew Eccles, Alan Lloyd Hodgkin and Andrew Fielding Huxley “for their discoveries concerning the ionic mechanisms involved in excitation and inhibition in the peripheral and central portions of the nerve cell membrane.”

– Nobel Prize Committee, 1963

By using voltage clamp experiments and varying extracellular sodium and potassium concentrations, Alan Hodgkin and Andrew Huxley described a model in 1952 to explain the ionic mechanisms underlying the initiation and propagation of action potentials. The Hodgkin-Huxley (HH) model is a set of four nonlinear ordinary differential equations that approximates the electrical characteristics of neurons firing. They received the 1963 Nobel Prize in Physiology or Medicine for their ground-breaking work on modeling neuron spikes. In this section, we will first review essential components of the HH model, includ-

ing the membrane capacitance, reversal potentials, active conductances, and membrane current. Then, we will present the mathematical framework of the HH model.

2.1 Membrane Capacitance and Reversal Potentials

Ionic pumps embedded in the membranes of nerve cells typically maintain a negative charge on the inside surface of the cell membrane, and a balancing positive charge on its outside surface. This charge imbalance creates a voltage difference V between the inside and outside of the cell. Specifically, the lipid bilayer of the cell membrane forms a thin insulator that separates two electrolytic media, the extracellular space and the cytoplasm [51]. The specific membrane capacitance C_m , the potential across the membrane V , and the amount of the excess charge density Q (per area) are related by the equation $Q = C_m V$ [1]. Given that the thickness of the membrane is a constant, the total membrane capacitance c_m of a cell is a quantity directly proportional to the membrane surface area and the properties of the membrane. Therefore, the total membrane capacitance is $c_m = C_m A$, where C_m is the per area membrane capacitance (typically in units of $\mu F/cm^2$) and A is the area (typically in units of cm^2). The capacitance per unit area of membrane, C_m , is approximately the same for all neurons with $C_m \approx 10nF/mm^2$.

The membrane capacitance can be used to determine the current required to change the membrane potential at a given rate. More specifically, the relation between the change in voltage and charge can be written as

$$C_m \frac{dV}{dt} = \frac{dQ}{dt}. \quad (2.1)$$

Equation 2.1 plays an important role in the formulation of the HH model.

The reversal potential for a channel is the voltage at which the net current through the channel is zero. The reversal potential is determined by the difference of concentration of ions inside the cell, $[C]_{\text{in}}$, and the concentration outside the cell, $[C]_{\text{out}}$. The Nernst equation [22] for the reversal potential can be written as

$$V = \frac{RT}{zF} \log \frac{[C]_{\text{in}}}{[C]_{\text{out}}}, \quad (2.2)$$

where R is the universal gas constant: $R = 8.31 \text{ JK}^{-1}\text{mol}^{-1}$, T is the temperature in Kelvins, F is the Faraday constant, and z is the charge of the ion species ($z = +1$ for Na^+ and K^+ , -1 for Cl^- , and $z = +2$ for Ca^{2+}). The reversal potential for a K^+ channel, V_{K} , typically falls in the range between -70 and -90 mV. The reversal potential for Na^+ , V_{Na} , is 50 mV or higher. Throughout this thesis, we will use $V_{\text{K}} = -77\text{mV}$ and $V_{\text{Na}} = 50\text{mV}$ for all numerical simulations.

2.2 The Membrane Current

The membrane current of a neuron is the total current flowing across the membrane through all ion channels [22]. The total membrane current is determined by including all currents resulting from different types of channels within the cell membrane. To make it comparable for neurons with different sizes, the membrane current per unit area of cell membrane is conveniently used, which we define as I_m . The total membrane current is obtained from the current per unit area I_m by multiplying by the total surface area of the cell. For each ion channel, the current depends approximately linearly on the difference between the membrane potential and its reversal potential, V_{ion} .¹ Given the conductance per unit area,

¹Here we followed the standard convention that equates “ion channels” with “ion”; in particular, we do not consider nonspecific ion channels.

g_{ion} , for each ion channel, the total membrane current can be expressed as a sum of currents from different ion channels as

$$I_m = \sum_{\text{ion}} g_{\text{ion}}(V - V_{\text{ion}}), \quad (2.3)$$

which is Ohmic approximation for the membrane current.²

2.3 The Hodgkin-Huxley Model

In the model Hodgkin and Huxley proposed in 1952, the behaviour of a nerve fiber is described using an electrical network where the membrane is represented by a capacitor of fixed capacitance, and the ion pathways through the membrane are represented by three resistance-capacitor modules arranged in parallel (see Fig. 2.1).

The Hodgkin-Huxley [59] model treats each component of an excitable cell as an electrical element (as shown in Figure 2.1). The lipid bilayer is represented as a capacitance (C_m). Voltage-gated ion channels are represented by active electrical conductances (g_{ion} , where ion stands for Na^+ or K^+) that vary dynamically, depending on the voltage. The leak channels (g_L) represent a generic ionic current. Reversal potentials for the Na^+ , K^+ and leak channels are defined as V_{Na} , V_{K} , and V_L , respectively.

Therefore, if we denote the membrane potential by V_m , the current flowing through the lipid bilayer can be written as

$$I_c = C_m \frac{dV_m}{dt} \quad (2.4)$$

²A more precise expression, given by the Goldman-Hodgkin-Katz equation, gives good agreement with this linear approximation over the physiologically relevant voltage range.

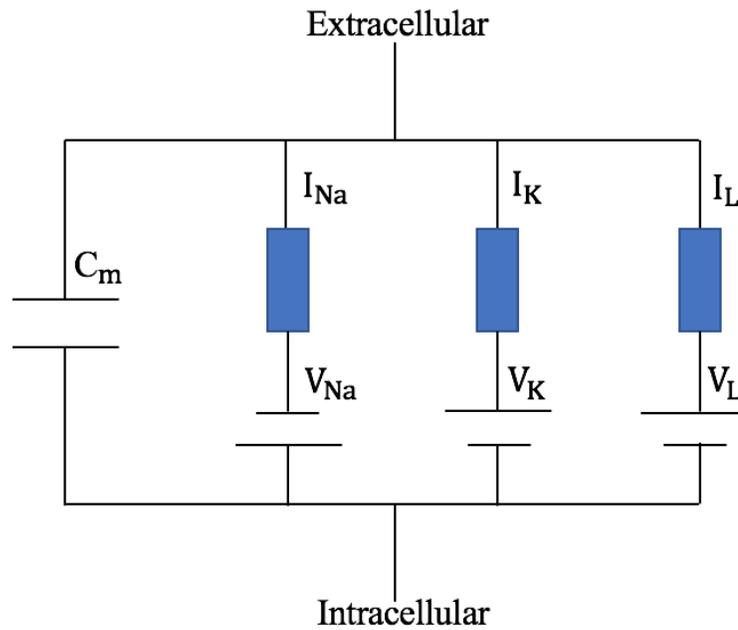


Figure 2.1: Basic components of the Hodgkin–Huxley model. The lipid bilayer is represented as a capacitance (C_m). Reversal potentials for the Na^+ , K^+ and leak channels are V_{Na} , V_{K} , and V_{L} , respectively. Shaded boxes represent active conductances (inverse resistances) g_{ion} that specify the Ohmic currents $I_{\text{ion}} = g_{\text{ion}}(V - V_{\text{ion}})$, for $\text{ion} \in \{\text{Na}, \text{K}, \text{L}\}$.

and the current through a given ion channel is a product

$$I_{\text{ion}} = g_{\text{ion}}(V_m - V_{\text{ion}}). \quad (2.5)$$

The original HH model only considered the Na^+ and K^+ currents and a leak current, therefore, the total current through the membrane is

$$I = C_m \frac{dV_m}{dt} + g_{\text{K}}(V_m - V_{\text{K}}) + g_{\text{Na}}(V_m - V_{\text{Na}}) + g_{\text{L}}(V_m - V_{\text{L}}) \quad (2.6)$$

where I is the total membrane current per unit area, C_m is the membrane capacitance per unit area, g_{K} and g_{Na} are the potassium and sodium conductances per unit area. V_{K} and V_{Na} are the potassium and sodium reversal potentials, and g_{L} and V_{L} are the leak conductance per unit area and leak reversal potential, respectively. Appendix B has a complete list of parameters.

Using a series of voltage clamp experiments, and by numerically fitting parameters, Hodgkin and Huxley [59] developed a set of four ordinary differential equations as

$$C \frac{dv}{dt} = -\bar{g}_{\text{Na}} m^3 h (v - V_{\text{Na}}) - \bar{g}_{\text{K}} n^4 (v - V_{\text{K}}) - g_{\text{L}} (v - V_{\text{L}}) + I_{\text{app}}, \quad (2.7)$$

$$\frac{dm}{dt} = \alpha_m(v)(1 - m) - \beta_m(v)m, \quad (2.8)$$

$$\frac{dh}{dt} = \alpha_h(v)(1 - h) - \beta_h(v)h, \quad (2.9)$$

$$\frac{dn}{dt} = \alpha_n(v)(1 - n) - \beta_n(v)n, \quad (2.10)$$

where v is the membrane potential, I_{app} is the applied current, and $0 \leq m, n, h \leq 1$ are dimensionless gating variables associated with the Na^+ and K^+ channels. The constant \bar{g}_{ion} is the maximal value of the conductance for the sodium and potassium channel, respectively.

Parameters V_{ion} and C are the ionic reversal potentials and capacitance, respectively. The quantities α_x and β_x , $x \in \{m, n, h\}$ are the voltage-dependent per capita transition rates, defined as

$$\alpha_m(v) = \frac{0.1(25 - v)}{\exp(2.5 - 0.1v) - 1}, \quad (2.11)$$

$$\beta_m(v) = 4 \exp(-v/18), \quad (2.12)$$

$$\alpha_h(v) = 0.07 \exp(-v/20), \quad (2.13)$$

$$\beta_h(v) = \frac{1}{\exp(3 - 0.1v) + 1}, \quad (2.14)$$

$$\alpha_n(v) = \frac{0.01(10 - v)}{\exp(1 - 0.1v) - 1}, \quad (2.15)$$

$$\beta_n(v) = 0.125 \exp(-v/80). \quad (2.16)$$

Fig. 2.2 shows the voltage component of a regular spiking trajectory of the HH equations with constant driving current injection of $I = 10 \text{ nA/cm}^2$.

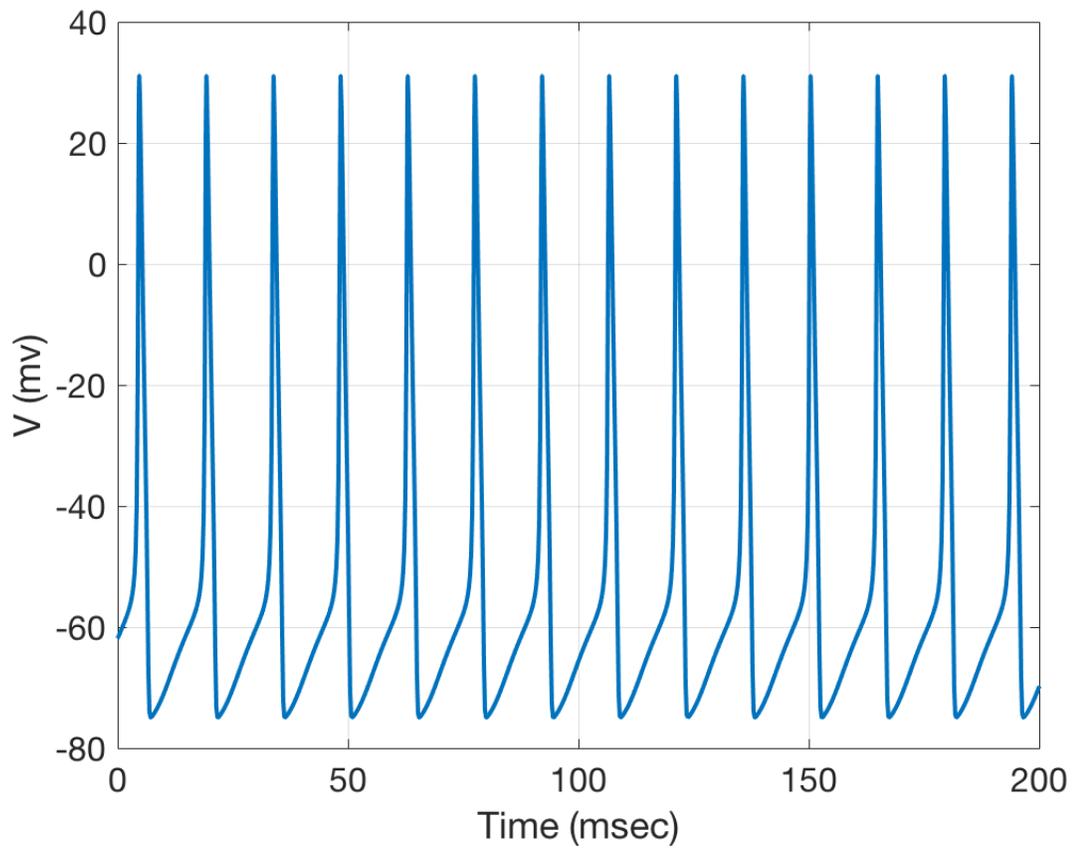


Figure 2.2: Sample trace of the deterministic HH model showing periodic firing using standard parameters (cf. Tab. B.1) and $I = 10 \text{ nA/cm}^2$. The interspike interval (ISI) for this specific case is $T \approx 14.6384 \text{ ms}$.

Chapter 3

Channel Noise

The probabilistic gating of voltage-dependent ion channels is a source of electrical “channel noise” in neurons. This noise has long been implicated in limiting the reliability of neuronal responses to repeated presentations of identical stimuli.

– White, Rubinstein and Kay [126]

Nerve cells communicate with one another, process sensory information, and control motor systems through transient voltage pulses, or spikes. At the single-cell level, neurons exhibit a combination of deterministic and stochastic behaviors. In the supra-threshold regime, the regular firing of action potentials under steady current drive suggests limit cycle dynamics, with the precise timing of voltage spikes perturbed by noise. Variability of action potential timing persists even under blockade of synaptic connections, consistent with an intrinsically noisy neural dynamics arising from the random gating of ion channel populations, or “channel noise” [126].

Channel noise arises from the random opening and closing of finite populations of ion channels embedded in the cell membranes of individual nerve cells, or localized regions of

Estimated numbers of Na ⁺ and K ⁺ channels in different cell types			
Ion	Type of cell	Number of channels	Reference
Na ⁺	chromaffin cells	1,800-12,500	[35, 120] ^a
	cardiac Purkinje cells	≥325,000	[79] ^b
	node of Ranvier	21,000-42,000	[115] ^c
	squid axon (1mm) ^d	≥18,800	[34] ^d
	pyramidal cell	≥17,000	[34] ^d
	Purkinje cell ^g	47,000-158,000	[38, 113] ^{d,f,g}
K ⁺	pre-BötC neurons ^h	56-5,600	[15, 34] ^{d,f,h}
	squid axon (1mm) ^d	≥5,600	[34] ^d
	pyramidal cell	≥2,000	[34] ^d
	Purkinje cell ^g	3,000-55,000	[38, 113] ^{d,e,g}
pre-BötC neurons ^h		112-2,240	[15, 34] ^{d,e,h}

Table 3.1: Details of the data sources:

- (a) Na⁺ density: 1.5-10 channels/ μm^2 [35], the average diameter of rounded chromaffin cells is $d \approx 20\mu\text{m}$, Area= πd^2 [120].
- (b) Na⁺ density: 260 channels/ μm^2 [79], and diameter of roughly $20\mu\text{m}$ [79].
- (c) Number of Na⁺ channels in Tab. 1 from [115].
- (d) Na⁺ density: 60 channels/ μm^2 in squid axon, and 68 channels/ μm^2 in pyramidal axon (Tab. S1 in [34]). K⁺ density: 18 channels/ μm^2 in squid axon, and 8 channels/ μm^2 in pyramidal axon (Tab. S1 in [34]). Membrane area: squid axon: $0.1\mu\text{m}$ diameter and 1mm length (Fig. S2 in [34]); pyramidal cell: $0.08\mu\text{m}$ diameter with 1mm length (Fig. S1 in [34]). Single voltage-gated ion channel conductance is typically in the range of 5-50 pS, and 15-25 pS for Na⁺ (p. 1148 [34]).
- (e) Single K⁺ channel conductance ([113]): inward rectifier in horizontal cells (20-30 pS in 62-125 mM K⁺, 9-14°C); skeletal muscle (10 pS in 155 mM K⁺, 24-26°C); egg cells (≈ 6 pS for 155 mM K⁺, 14-15°C); heart cells (27 pS for 145 mM K⁺ at 17-23°C; 45 pS for 150 mM K⁺ at 31-36°C).
- (f) Single Na⁺ channel conductance is ≈ 14 pS in squid axon, other measurements under various conditions show results in the range of 2-18 pS (Tab. 1 in [9]).
- (g) Maximal conductance for different K⁺ channels (Tab. 1 in [38]): SK K⁺ (10 mS/ cm^2), highly TEA K⁺ (41.6 mS/ cm^2) sensitive BK K⁺ (72.8 mS/ cm^2); membrane area ($1521\mu\text{m}^2$). Maximal conductance for resurgent Na⁺ (156 mS/ cm^2). Note that the range of K⁺ channels provided here is for each type of K⁺ channel, not the total number of K⁺ channels.
- (h) Maximal conductance (\bar{g}_{ion}) in pacemaker cells of the pre-Bötzinger complex (pre-BötC) [15]: $\bar{g}_{\text{NaP}} = 2.8$ nS for persistent Na⁺ current, $\bar{g}_{\text{Na}} = 28$ nS for fast Na⁺ current, and $\bar{g}_{\text{K}} \in [5.6, 11.2]$ nS for different types of K⁺ channels (p. 384-385).

axons or dendrites. Electrophysiological and neuroanatomical measurements do not typically provide direct measures of the sizes of ion channel populations. Rather, the size of ion channel populations must be inferred indirectly from other measurements. Several papers report the density of sodium or potassium channels per area of cell membrane [34, 35, 79]. Multiplying such a density by an estimate of the total membrane area of a cell gives one estimate for the size of a population of ion channels. Sigworth [115] pioneered statistical measures of ion channel populations based on the mean and variance of current fluctuations observed in excitable membranes, for instance in the isolated node of Ranvier in axons of the frog. Single-channel recordings [85] allowed direct measurement of the “unitary”, or single-channel-conductance, g_{Na}^o or g_{K}^o . Most conductance-based, ordinary differential equations models of neural dynamics incorporate maximal conductance parameters (\bar{g}_{Na} or \bar{g}_{K}) which nominally represents the conductance that would be present if all channels of a given type were open. The ratio of \bar{g} to g^o thus gives an indirect estimate of the number of ion channels in a specific cell type. Tab. 3.1 summarizes a range of estimates for ion channel populations from several sources in the literature. Individual cells range from 50 to 325,000 channels for each type of ion. In §11.2 of this thesis, we will consider effective channel populations spanning this entire range (cf. Figs 11.3, 11.4).

3.1 Modeling Channel Noise

Hodgkin and Huxley’s quantitative model for active sodium and potassium currents producing action potential generation in the giant axon of *Loligo* [59] suggested an underlying system of gating variables consistent with a multi-state Markov process description [58]. The discrete nature of individual ion channel conductances was subsequently experimentally confirmed [85]. Following this work, numerical studies of high-dimensional discrete-state,

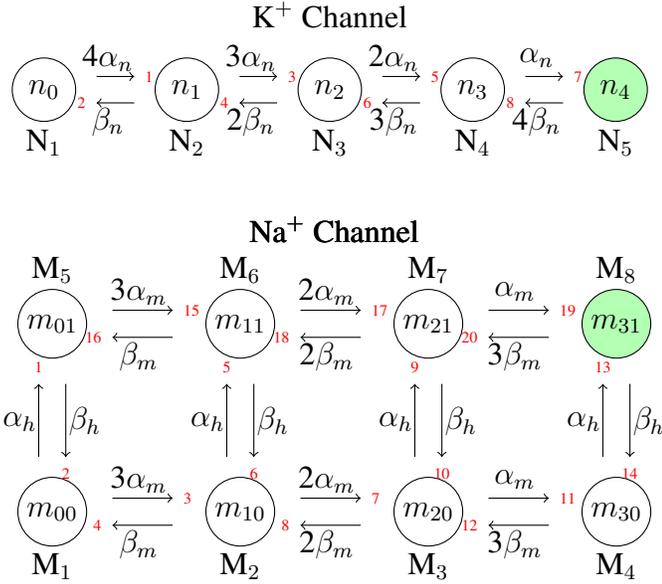


Figure 3.1: Molecular potassium (K⁺) and sodium (Na⁺) channel states for the Hodgkin-Huxley model. Filled circles mark conducting states n_4 and m_{31} . Per capita transition rates for each directed edge (α_n , β_n , α_m , β_m , α_h and β_h) are voltage dependent (cf. eqns. (B.1)-(B.6)). Directed edges are numbered 1-8 (K⁺ channel) and 1-20 (Na⁺-channel), marked in small red numerals.

continuous-time Markov chain models produced insights into the effects of fluctuations in discrete ion channel populations on action potentials [116, 119], aka *channel noise* [125, 126].

In the standard molecular-level HH model, which we adopt here, the K⁺ channel comprises four identical “ n ” gates that open and close independently, giving a five-vertex channel-state diagram with eight directed edges; the channel conducts a current only when in the rightmost state (Fig. 3.1, top). The Na⁺ channel comprises three identical “ m ” gates and a single “ h ” gate, all independent, giving an eight-vertex diagram with twenty directed edges, of which one is conducting (Fig. 3.1, bottom).

Discrete-state channel noise models are numerically intensive, whether implemented using discrete-time binomial approximations to the underlying continuous-time Markov

process [102, 116] or continuous-time hybrid Markov models with exponentially distributed state transitions and continuously varying membrane potential. The latter were introduced by [18] and are in principle exact [4]. Under voltage-clamp conditions the hybrid conductance-based model reduces to a time-homogeneous Markov chain [19] that can be simulated using standard methods such as Gillespie’s exact algorithm [46, 47]. Even with this simplification, such Markov Chain (MC) algorithms are numerically expensive to simulate with realistic population sizes of thousands of channels or greater. Therefore, there is an ongoing need for efficient and accurate approximation methods.

Following Clay and DeFelice’s exposition of continuous time Markov chain implementations, [39] introduced a Fokker-Planck equation (FPE) framework that captured the first and second order statistics of HH ion channel dynamics in a 14-dimensional representation. Taking into account conservation of probability, one needs four variables to represent the population of K^+ channels, seven for Na^+ , and one for voltage, leading to a 12-dimensional state space description. The resulting high-dimensional partial differential equation is impractical to solve numerically. However, as Fox and Lu observed, “to every Fokker-Planck description, there is associated a Langevin description” [39]. They therefore introduced a Langevin stochastic differential equation of the form:

$$C \frac{dV}{dt} = I_{\text{app}}(t) - \bar{g}_{\text{Na}} \mathbf{M}_8 (V - V_{\text{Na}}) - \bar{g}_{\text{K}} \mathbf{N}_5 (V - V_{\text{K}}) - g_{\text{leak}} (V - V_{\text{leak}}), \quad (3.1)$$

$$\frac{d\mathbf{M}}{dt} = A_{\text{Na}} \mathbf{M} + S_1 \xi_1, \quad (3.2)$$

$$\frac{d\mathbf{N}}{dt} = A_{\text{K}} \mathbf{N} + S_2 \xi_2, \quad (3.3)$$

where C is the capacitance, I_{app} is the applied current, maximal conductances are denoted \bar{g}_{ion} , with V_{ion} being the associated reversal potential, and ohmic leak current $g_{\text{leak}}(V - V_{\text{leak}})$. $\mathbf{M} \in \mathbb{R}^8$ and $\mathbf{N} \in \mathbb{R}^5$ are vectors for the fractions of Na^+ and K^+ channels in each

state, with M_8 representing the open channel fraction for Na^+ , and N_5 the open channel fraction for K^+ (Fig. 3.1). Vectors $\xi_1(t) \in \mathbb{R}^8$ and $\xi_2(t) \in \mathbb{R}^5$ are independent Gaussian white noise processes with zero mean and unit variances $\langle \xi_1(t) \xi_1^T(t') \rangle = I_8 \delta(t - t')$ and $\langle \xi_2(t) \xi_2^T(t') \rangle = I_5 \delta(t - t')$. The state-dependent rate matrices A_{Na} and A_{K} are given in eqns. (5.10) and (5.11). In Fox and Lu’s formulation, S must satisfy $S = \sqrt{D}$, where D is a symmetric, positive semi-definite $k \times k$ “diffusion matrix” (see Appendix C for the D matrices for the standard HH K^+ and Na^+ channels). We will refer to the 14-dimensional Langevin equations (3.1)-(3.3), with $S = \sqrt{D}$, as the “Fox-Lu” model.

3.2 Motivation 1: The Need for Efficient Models

The original Fox-Lu model, later called the “conductance noise model” by [49], did not see widespread use until gains in computing speed made the square root calculations more feasible. Seeking a more efficient approximation, [39] also introduced a four-dimensional Langevin version of the HH model. This model was systematically studied in [40] which can be written as follows:

$$C \frac{dV}{dt} = I_{\text{app}}(t) - \bar{g}_{\text{Na}} m^3 h (V - V_{\text{Na}}) - \bar{g}_{\text{K}} n^4 (V - V_{\text{K}}) - g_{\text{leak}} (V - V_{\text{leak}}) \quad (3.4)$$

$$\frac{dx}{dt} = \alpha_x (1 - x) - \beta_x x + \xi_x(t), \text{ where } x = m, h, \text{ or, } n. \quad (3.5)$$

where $\xi_x(t)$ are Gaussian processes with covariance function

$$E[\xi_x(t), \xi_x(t')] = \frac{\alpha_x (1 - x) + \beta_x x}{N} \delta(t - t'). \quad (3.6)$$

Here N represents the total number of Na^+ channels (respectively, the total number of K^+ channels) and $\delta(\cdot)$ is the Dirac delta function. This model, referred as the “subunit noise model” by [49], has been widely used as an approximation to MC ion channel models (see references in [12, 49]). For example, [103] used this approximation to investigate stochastic resonance and coherence resonance in forced and unforced versions of the HH model (e.g. in the excitable regime). However, the numerical accuracy of this method was criticized by several studies [12, 81], which found that its accuracy does not improve even with increasing numbers of channels.

Although more accurate approximations based on Gillespie’s algorithm (using a piecewise constant propensity approximation, [12, 81]) and even based on exact simulations [4, 18, 87] became available, they remained prohibitively expensive for large network simulations. Meanwhile, Goldwyn and Shea-Brown’s rediscovery of Fox and Lu’s earlier conductance based model [48, 49] launched a flurry of activity seeking the best Langevin-type approximation. The paper [49] introduced a faster decomposition algorithm to simulate equations (3.1)-(3.3), and showed that Fox and Lu’s method accurately captured the fractions of open channels and the inter-spike interval (ISI) statistics, in comparison with Gillespie-type Monte Carlo (MC) simulations. However, despite the development of efficient singular value decomposition based algorithms for solving $S = \sqrt{D}$, this step still causes a bottleneck in the algorithms based on [39, 48, 49].

The persistent need for fast and accurate simulation methods is the first main motivation of this thesis work. Many variations on Fox and Lu’s 1994 Langevin model have been proposed in recent years [20, 21, 41, 56, 60, 61, 73, 90, 93] including Goldwyn *et al*’s work [48, 49], each with its own strengths and weaknesses. One class of methods imposes projected boundary conditions [20, 21]; as we will show in §8, this approach leads to an inaccurate interspike interval distribution, and is inconsistent with a natural multinomial invariant

manifold structure for the ion channels. Several methods implement correlated noise at the subunit level, as in (3.5)-(3.6) [40, 55, 56, 73]. In the subunit model (cf. eqn. (3.5)) a single noise source represents the fluctuations associated with the gating variable x ($x \in \{m, n, h\}$). However, if one recognizes that, at the molecular level, the *individual directed edges* in Fig. 3.1 represent the independent noise sources in ion channel dynamics, then it becomes clear that the subunit noise model obscures the biophysical origin of ion channel fluctuations. Some methods introduce the noisy dynamics at the level of edges rather than nodes, but lump reciprocal edges together into pairs [21, 60, 90, 93]. This approach implicitly assumes, in effect, that the ion channel probability distribution satisfies a detailed balance (or microscopic reversibility) condition. However, while detailed balance holds for the HH model under stationary voltage clamp, this condition is violated during active spiking. Finally, the stochastic shielding approximation [102, 104, 105] does not have a natural formulation in the representation associated with an $n \times n$ noise coefficient matrix S ; in the cases of rectangular S matrices used in [21, 90] stochastic shielding can only be applied to reciprocal pairs of edges. We will elaborate on these points in §12.

Chapter 4

The Connection Between Variance of ISIs with the Random Gating of Ion Channels

4.1 Variability In Action Potentials

At the single-cell level, neurons exhibit a combination of deterministic and stochastic behaviors. In the supra-threshold regime, the action potentials under steady current drive suggests limit cycle dynamics, with the precise timing of voltage spikes perturbed by noise. Variability of action potential timing persists even under blockade of synaptic connections, consistent with an intrinsically noisy neural dynamics arising from the random gating of ion channel populations, or “channel noise” [126]. Channel noise can have a significant effect on spike generation [78, 107], propagation along axons [33], and spontaneous (ectopic) action potential generation in the absence of stimulation [88]. At the network level, channel

noise can drive endogenous variability of vital rhythms such as respiratory activity [134].

Understanding the molecular origins of spike time variability may shed light on several phenomena in which channel noise plays a role. For example, microscopic noise can give rise to a stochastic resonance behavior [103], and can contribute to cellular- and systems-level timing changes in the integrative properties of neurons [26]. Jitter in spike times under steady drive may be observed in neurons from the auditory system [45, 50, 83] as well as in the cerebral cortex [78] and may play a role in both fidelity of sensory information processing and in precision of motor control [107].

4.2 Motivation 2: Understanding the Molecular Origins of Spike Time Variability

As a motivating example for this dissertation, channel noise is thought to underlie jitter in spike timing observed in cerebellar Purkinje cells recorded in vitro from the “leaner mouse”, a P/Q-type calcium channel mutant with profound ataxia [123]. Purkinje cells fire Na^+ action potentials spontaneously [76, 77], and may do so at a very regular rate [123], even in the absence of synaptic input (cf. Fig. 4.1 A and C). Mutations in an homologous human calcium channel gene are associated with episodic ataxia type II, a debilitating form of dyskinesia [94, 97]. Previous work has shown that the leaner mutation increases the variability of spontaneous action potential firing in Purkinje cells [91, 123] (Fig. 4.1 B and D). It has been proposed that increased channel noise akin to that observed in the leaner mutant plays a mechanistic role in this human disease [123].

To understand the underlying mechanisms for the different firing properties in wild type and the leaner mutant, a model with biological fidelity is desired. For example, in addition to fast channel noise due to sodium and potassium channels, fluctuations of electrical

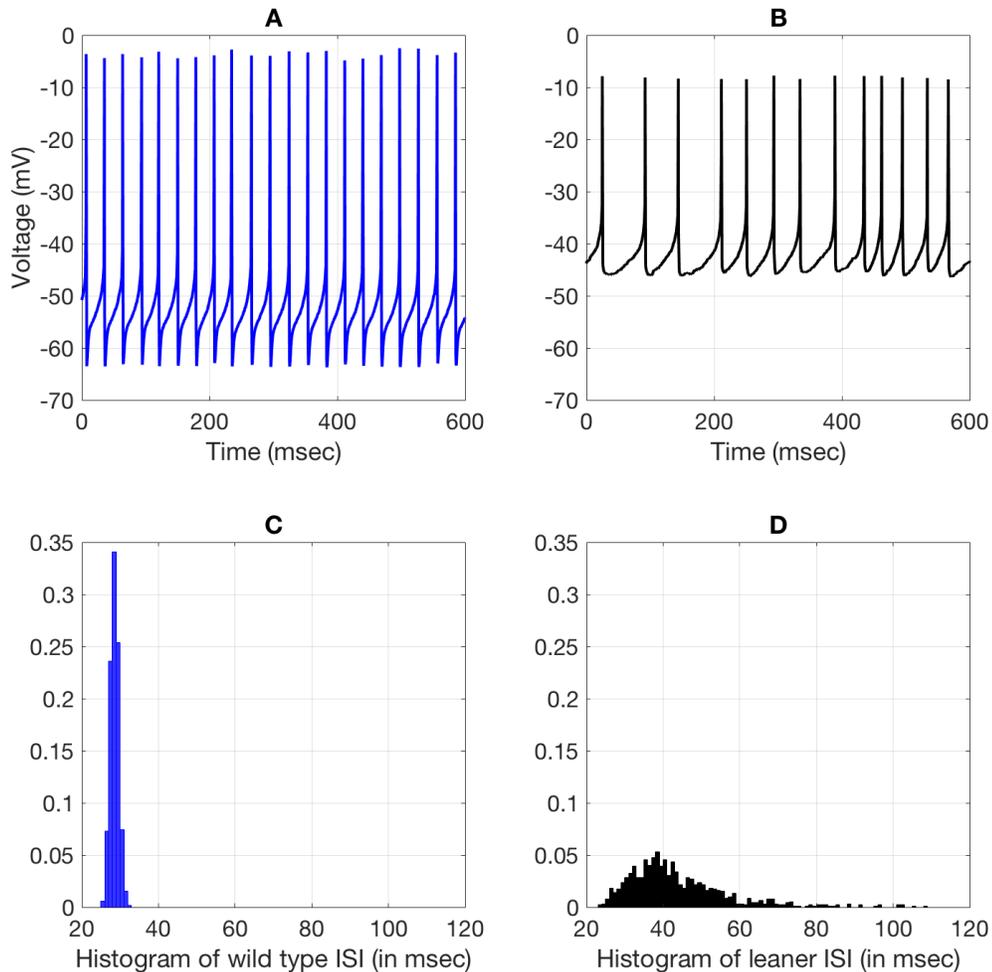


Figure 4.1: Somatic voltage recorded in vitro from intact Purkinje cells (cerebellar slice preparation) during spontaneous tonic firing, with synaptic input blocked. Voltage was sampled at 20KHz and digitized at 32/mV. Recordings performed in the laboratory of Dr. David Friel (CWRU School of Medicine). **A**: Sample voltage recordings from a wild type Purkinje cell showing precise spontaneous firing with interspike interval (ISI) coefficient of variation ($CV = \text{standard deviation} / \text{mean ISI}$) of approximately 3.9%. **B**: Sample recordings from Purkinje cells with leaner mutation in P/Q-type calcium channels showing twice the variability in ISI ($CV \approx 30\%$). **C**, **D**: histogram of ISI for wild type and leaner mutation, respectively. Bin width = 1 msec for each.

activity in PCs may be subject to the effects of slow noise processes such as stochasticity of calcium channels, calcium-gated potassium channels, and dendritic filtering. Cerebellar Purkinje cells have been studied using models with a wide range of complexity, from models with thousands of subcellular compartments each with multiple gating variables and voltage [23, 24, 25] to “reduced” models with only dozens of compartments [38] as well as models with as few as five dynamical variables [36]. The currents at work in Purkinje cells have also been subject to detailed modeling, including “resurgent” sodium current [17, 98], multiple types of potassium currents [37, 66], calcium currents [37, 82] and calcium-dependent potassium currents [37]. In order to pave the way for tackling more complex models, in this thesis we restrict attention to a simpler, single-compartment conductance-based model, the canonical excitable membrane model originating with Hodgkin and Huxley [59].

Despite its practical importance, a quantitative understanding of distinct molecular sources of macroscopic timing variability remains elusive, even for the HH model. As the second main motivation of this thesis, we would like to study the quantitative connection between the molecular-level ion channel transitions to the macroscopic variability of timing in membrane action potentials. Significant theoretical attention has been paid to the variance of phase response curves and interspike interval (ISI) variability. Most analytical studies are based on the integrate-and-fire model [13, 74, 122], except [31], which perturbs the voltage of a conductance-based model with a white noise current rather than through biophysically-based channel noise. Standard models of stochastic ion channel kinetics comprise hybrid stochastic systems. As illustrated in Fig. 4.2, the membrane potential evolves deterministically, given the transmembrane currents; the currents are determined by the ion channel state; the ion channel states fluctuate stochastically with opening and closing rates that depend on the voltage [4, 10, 14, 92]. This closed-loop

nonlinear dynamical stochastic system is difficult to study analytically, because of recurrent statistical dependencies of the variables with one another. An important and well studied special case is fixed-voltage clamp, which reduces the ion channel subsystem to a time invariant Markov process [105]. Under the more natural current clamp, the ion channel dynamics taken alone are no longer Markovian, as they intertwine with current and voltage. A priori, it is challenging to draw a direct connection between the variability of spike timing and molecular-level stochastic events, such as the opening and closing of specific ion channels, as spike timing is a pathwise property reflecting the effects of fluctuations accumulated around each orbit or cycle.

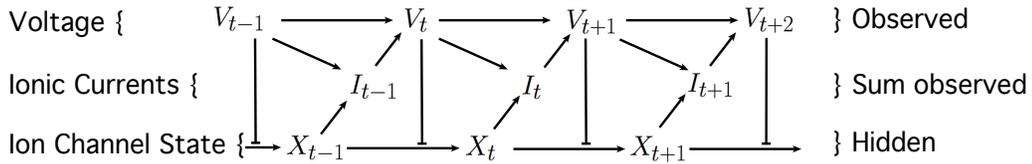


Figure 4.2: Interdependent variables under current clamp. Statistical dependencies among voltage (V_t), ionic currents (I_t), and ion channel state (X_t) form a hybrid, or piecewise deterministic, stochastic model. Voltage V_t at time t influences current I_t as well as the transition from channel state $X = [M_{00}, M_{10}, \dots, M_{31}, N_0, N_1, \dots, N_4]$, at time $t - 1$ to time t . (For illustration we assume discrete sampling at integer multiples of a nominal time step Δt in arbitrary units.) Channel state dictates the subsequent ionic current, which dictates the voltage increment. Arrowheads (\rightarrow) denote deterministic dependencies. T-connectives (\perp) denote statistical dependencies.

In [102] Schmandt and Galán introduced *stochastic shielding* as a fast, accurate approximation scheme for stochastic ion channel models. Rather than simplifying the Markov process by aggregating ion channel states, stochastic shielding reduces the complexity of the underlying sample space by eliminating independent noise sources (corresponding to individual directed edges in the channel state transition graph) that make minimal contributions to ion channel state fluctuations. In addition to providing an efficient numerical procedure, stochastic shielding leads to an *edge importance measure* [105] that quantifies

the contribution of the fluctuations arising along each directed edge to the variance of channel state occupancy (and hence the variance of the transmembrane current). The stochastic shielding method then amounts to simulating a stochastic conductance-based model using only the noise terms from the most important transitions. While the original, heuristic implementation of stochastic shielding considered both current and voltage clamp scenarios [102], subsequent mathematical analysis of stochastic shielding considered only the constant voltage-clamp case [104, 105]. In our recent paper [96] we provide, to our knowledge, the first analytical treatment of the variability of spike timing under current clamp arising from the random gating of ion channels with realistic (Hodgkin-Huxley) kinetics. Building on prior work [75, 95, 102, 104, 105], we study the variance of the transition times among a family of Poincaré sections, the mean–return-time (MRT) isochrons investigated by [75, 108] that extend the notion of phase reduction to stochastic limit cycle oscillators. We prove a theorem that gives the form of the variance, σ_ϕ^2 , of inter-phase-intervals (IPI)¹ in the limit of small noise (equivalently, large channel number or system size), as a sum of contributions $\sigma_{\phi,k}^2$ from each directed edge k in the ion channel state transition graph (Fig. 3.1). The IPI variability involves several quantities: the per capita transition rates α_k along each transition, the mean-field ion channel population $M_{i(k)}$ at the source node for each transition, the stoichiometry (state-change) vector ζ_k for the k th transition, and the phase response curve \mathbf{Z} of the underlying limit cycle:

$$\sigma_\phi^2 = \sum_{k \in \text{all edges}} \sigma_{\phi,k}^2 = \epsilon \bar{T}_0 \sum_k \mathbb{E} \left(\alpha_k(v(t)) M_{i(k)}(t) (\zeta_k^\top \mathbf{Z}(t))^2 dt \right) + O(\epsilon^2),$$

in the limit as $\epsilon \rightarrow 0^+$. Here \bar{T}_0 , $v(t)$ and $\mathbf{M}(t)$ are the period, voltage, and ion channel population vector of the deterministic limit cycle for $\epsilon = 0$. \mathbb{E} denotes expectation with

¹Equivalently, “iso-phase-intervals”: the time taken to complete one full oscillation, from a given isochron back to the same isochron.

respect to the stationary probability density for the Langiven model (cf. eqn. (9.28)). As detailed below, we scale $\epsilon \propto 1/\sqrt{\Omega}$ where the system size Ω reflects the size of the underlying ion channel populations.

Thus we are able to pull apart the distinct contribution of each independent source of noise (each directed edge in the ion channel state transition graphs) to the variability of timing. Figs. 11.4-11.5 illustrate the additivity of contributions from separate edges for small noise levels. As a consequence of this linear decomposition, we can extend the stochastic shielding approximation, introduced in [102] and rigorously analyzed under voltage clamp in [104, 105], to the current clamp case. Our theoretical result guarantees that, for small noise, we can replace a full stochastic simulation with a more efficient simulation driven by noise from only the most “important” transitions with negligible loss of accuracy. We find numerically that the range of validity of the stochastic shielding approximation under current clamp extends beyond the “small noise limit” to include physiologically relevant population sizes, cf. Fig. 11.5.

The inter-phase-interval (IPI) is a mathematical construct closely related to, but distinct from, the inter-spike-interval (ISI). The ISI, determined by the times at which the cell voltage moves upward (say) through a prescribed voltage threshold v_{thresh} , is directly observable from experimental recordings – unlike the IPI. However, we note that both in experimental data and in stochastic numerical simulations, the variance of the ISI is not insensitive to the choice of voltage threshold, but increases monotonically as a function of v_{thresh} (cf. Fig. 11.2). In contrast, the variance of inter-*phase*-interval times is the same, regardless of which MRT isochron is used to define the intervals. This invariance property gives additional motivation for investigating the variance of the IPI.

The thesis is organized as follows. In Part I, we have presented an overview of the background information for single cell neurophysiology and also offered motivations both

from experimental recordings and from a theoretical point of view.

In Part II, we systematically study the canonical *deterministic* 14D version of the HH model. We prove a series of lemmas which show that (1) the multinomial submanifold \mathcal{M} is an invariant manifold within the 14D space and (2) the velocity on the 14D space and the pushforward of the velocity on the 4D space are identical. Moreover, we show (numerically) that (3) the submanifold \mathcal{M} is globally attracting, even under current clamp conditions. In §6, we describe our 14×28 Langevin HH model. Like [21, 90, 93], we avoid matrix decomposition by computing the coefficient matrix S directly. The key difference between our approach and its closest relative [93] is to use a rectangular $n \times k$ matrix S for which *each directed edge* is treated as an independent noise source, rather than lumping reciprocal edges together in pairs. In the new Langevin model, the form of our S matrix reflects the biophysical origins of the underlying channel noise, and allows us to apply the stochastic shielding approximation by neglecting the noise on selected individual directed edges.

In Part III, we answer an open question in the literature, arising from the fact that the decomposition $D = SS^T$ is not unique. As Prof. Fox has pointed out, sub-block determinants of the D matrices play a major role in the structure of the S matrix elements. In [41] this author conjectured that “a universal form for S may exist”. We obtained the universal form for the noise coefficient matrix S in [95], which we will review below in §7. Moreover, we prove that our model is equivalent to Fox and Lu’s 1994 model in the strong sense of *pathwise equivalence*. As we establish in §7, our model (without the stochastic shielding approximation) is pathwise equivalent to all those in a particular class of biophysically derived Langevin models, including those used in [39, 41, 48, 49, 90, 93]. In §8, we compare our Langevin model to several alternative stochastic neural models in terms of accuracy (of the full ISI distribution) and numerical efficiency.

In Part IV, we prove a theorem that decomposes the macroscopic variance of iso-phase intervals (IPIs) as a sum of contributions from the microscopic ion channel transitions in the limit of small noise. In §9, we state definitions, notations and terminology that are necessary for the proof of the theorem. We provide a detailed prediction of contributions to the variance of IPIs from each individual edges in Fig. 3.1. In §11, we test the numerical performance of the decomposition theorem and generalize it to the variance of inter-spike intervals.

In Part V, we conclude the thesis and discuss related work, as well as some limitations of our results.

Part II

Mathematical Framework

Chapter 5

The Deterministic 4-D and 14-D HH Models

In this chapter, we briefly review the classical four-dimensional model of [59] (HH), as well as its natural fourteen-dimensional version ([22], §5.7), with variables comprising membrane voltage and the occupancies of five potassium channel states and eight sodium channel states. The deterministic 14D model is the mean field of the channel-based Langevin model proposed by [39]; this thesis describes both the Langevin and the mean field versions of the 14D Hodgkin-Huxley system. For completeness of exposition, we briefly review the 4D deterministic HH system and its 14D deterministic counterpart. In §7 we will prove that the sample paths of a class of Langevin stochastic HH models are equivalent; in §5.3 we review analogous results relating trajectories of the 4D and 14D deterministic ODE systems.

In particular, we will show that the deterministic 14D model and the original 4D HH model are dynamically equivalent, in the sense that every flow (solution) of the 4D model

corresponds to a flow of the 14D model. The consistency of trajectories between of the 14D and 4D models is easy to verify for initial data on a 4D submanifold of the 14D space given by choosing multinomial distributions for the gating variables [22, 48]. Similarly, Keener established results on multinomial distributions as invariant submanifolds of Markov models with ion channel kinetics under several circumstances [29, 30, 63, 64], but without treating the general current-clamped case. Consistent with these results, we show below that the set of all 4D flows maps to an invariant submanifold of the state space of the 14D model. Moreover, we show numerically that solutions of the 14D model with arbitrary initial conditions converge to this submanifold. Thus the original HH model “lives inside” the 14D deterministic model in the sense that the former embeds naturally and consistently within the latter (cf. Fig. 5.1).

In the stochastic case, the 14D model has a natural interpretation as a hybrid stochastic system with independent noise forcing along each edge of the potassium (8 directed edges) and sodium (20 directed edges) channel state transition graphs. The hybrid model leads naturally to a biophysically grounded Langevin model that we describe in section §6. In contrast to the ODE case, the stochastic versions of the 4D and 14D models are *not* equivalent [49].

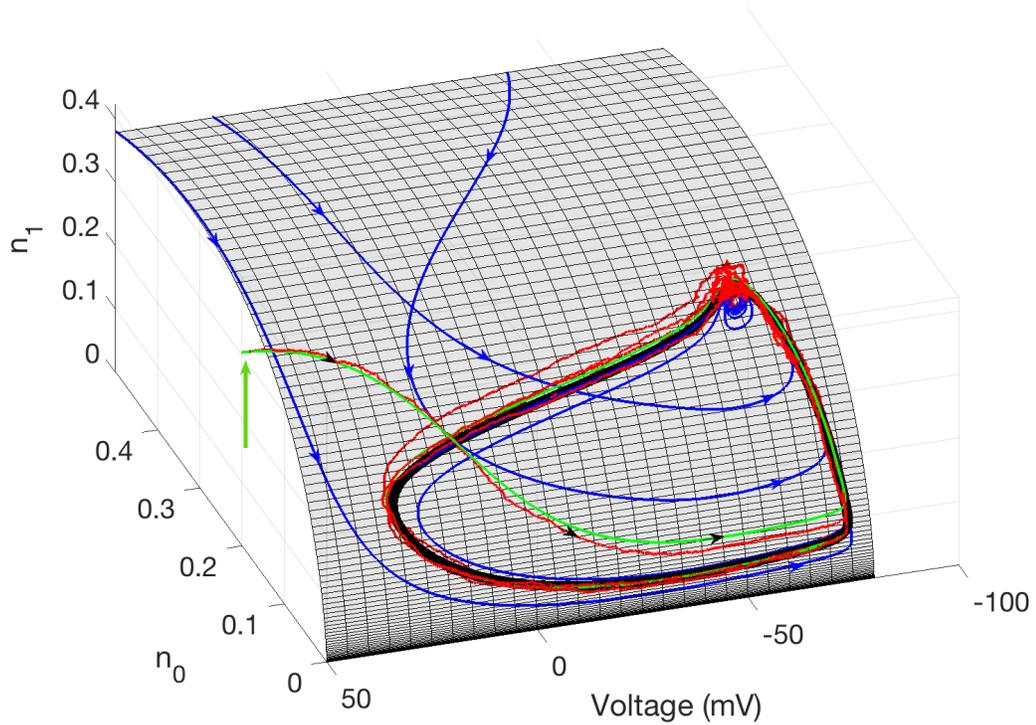


Figure 5.1: 4D and 14D HH models. The meshed surface is a three dimensional projection of the 14D state space onto three axes representing the voltage, v , the probability of all four potassium gates being in the closed state, n_0 , and the probability of exactly one potassium gate being in the open state, n_1 . **Blue curves:** Trajectories of the deterministic 14D HH model with initial conditions located on the 4D multinomial invariant submanifold, \mathcal{M} . We prove that \mathcal{M} is an invariant submanifold in §5.3. **Black curve:** The deterministic limit cycle solution for the 14D HH model, which forms a closed loop within \mathcal{M} . **Green curve:** A trajectory of the deterministic 14D HH model with initial conditions (vertical green arrow) off the multinomial submanifold. **Red curve:** A trajectory of the stochastic 14D HH model (cf. §6) with the same initial conditions as the green trajectory. The blue and black arrows mark the directions of the trajectories. Note that trajectories starting away from \mathcal{M} converge to \mathcal{M} ; and *all* deterministic trajectories converge to the deterministic limit cycle. Parameters of the simulation are given in Tab. B.1.

5.1 The 4D Hodgkin-Huxley Model

Recall that the 4D voltage-gated ion channel HH model is given by

$$C \frac{dv}{dt} = -\bar{g}_{\text{Na}} m^3 h (v - V_{\text{Na}}) - \bar{g}_{\text{K}} n^4 (v - V_{\text{K}}) - g_{\text{L}} (v - V_{\text{L}}) + I_{\text{app}}, \quad (5.1)$$

$$\frac{dm}{dt} = \alpha_m(v)(1 - m) - \beta_m(v)m, \quad (5.2)$$

$$\frac{dh}{dt} = \alpha_h(v)(1 - h) - \beta_h(v)h, \quad (5.3)$$

$$\frac{dn}{dt} = \alpha_n(v)(1 - n) - \beta_n(v)n, \quad (5.4)$$

where all variables and parameters retain the same meaning as in §2.3 (listed in App. B).

This system is a C^∞ vector field on a four-dimensional manifold (with boundary) contained in \mathbb{R}^4 : $\mathcal{X} = \{-\infty < v < \infty, 0 \leq m, h, n \leq 1\} = \mathbb{R} \times [0, 1]^3$. The manifold is forward and backward invariant in time. If I_{app} is constant then \mathcal{X} has an invariant subset given by $\mathcal{X} \cap \{v_{\min} \leq v \leq v_{\max}\}$, where v_{\min} and v_{\max} are calculated in Lemma 1.

As pointed out in ([65], §3, p. 106) and [63], for voltage either fixed or given as a prescribed function of time, the equations for m, h and n can be interpreted as the parametrization of an invariant manifold embedded in a higher-dimensional time-varying Markov system. Several papers developed this idea for a variety of ion channel models and related systems [30, 63] but the theory developed is restricted to the voltage-clamped case.

Under fixed voltage clamp, the ion channels form a time-homogeneous Markov process with a unique (voltage-dependent) stationary probability distribution. Under *time-varying* current clamp the ion channels nevertheless form a Markov process, albeit no longer time-homogeneous. Under these conditions the ion channel state converges rapidly to a multinomial distribution indexed by a low-dimensional set of time-varying parameters $(m(t), h(t), n(t))$ [64]. In the current-clamped case, the ion channel process, considered

alone, is neither stationary nor Markovian, making the analysis of this case significantly more challenging, from a mathematical point of view.

5.2 The Deterministic 14D Hodgkin-Huxley Model

For the HH kinetics given in Fig. 3.1 (on page 17), we define the eight-component state vector \mathbf{M} for the Na^+ gates, and the five-component state vector \mathbf{N} for the K^+ gates, respectively, as

$$\mathbf{M} = [m_{00}, m_{10}, m_{20}, m_{30}, m_{01}, m_{11}, m_{21}, m_{31}]^\top \in [0, 1]^8 \quad (5.5)$$

$$\mathbf{N} = [n_0, n_1, n_2, n_3, n_4]^\top \in [0, 1]^5, \quad (5.6)$$

where $\sum_{i=0}^3 \sum_{j=0}^1 m_{ij} = 1$ and $\sum_{i=0}^4 n_i = 1$. The open probability for the Na^+ channel is $\mathbf{M}_8 = m_{31}$, and is $\mathbf{N}_5 = n_4$ for the K^+ channel. The deterministic 14D HH equations may be written (compare (5.1)-(5.4))

$$C \frac{dV}{dt} = -\bar{g}_{\text{Na}} \mathbf{M}_8 (V - V_{\text{Na}}) - \bar{g}_{\text{K}} \mathbf{N}_5 (V - V_{\text{K}}) - g_{\text{L}} (V - V_{\text{L}}) + I_{\text{app}}, \quad (5.7)$$

$$\frac{d\mathbf{M}}{dt} = A_{\text{Na}}(V) \mathbf{M}, \quad (5.8)$$

$$\frac{d\mathbf{N}}{dt} = A_{\text{K}}(V) \mathbf{N}, \quad (5.9)$$

where the voltage-dependent drift matrices A_{Na} and A_{K} are given by

$$A_{\text{Na}}(V) = \begin{bmatrix} A_{\text{Na}}(1) & \beta_m & 0 & 0 & \beta_h & 0 & 0 & 0 \\ 3\alpha_m & A_{\text{Na}}(2) & 2\beta_m & 0 & 0 & \beta_h & 0 & 0 \\ 0 & 2\alpha_m & A_{\text{Na}}(3) & 3\beta_m & 0 & 0 & \beta_h & 0 \\ 0 & 0 & \alpha_m & A_{\text{Na}}(4) & 0 & 0 & 0 & \beta_h \\ \alpha_h & 0 & 0 & 0 & A_{\text{Na}}(5) & \beta_m & 0 & 0 \\ 0 & \alpha_h & 0 & 0 & 3\alpha_m & A_{\text{Na}}(6) & 2\beta_m & 0 \\ 0 & 0 & \alpha_h & 0 & 0 & 2\alpha_m & A_{\text{Na}}(7) & 3\beta_m \\ 0 & 0 & 0 & \alpha_h & 0 & 0 & \alpha_m & A_{\text{Na}}(8) \end{bmatrix}, \quad (5.10)$$

$$A_{\text{K}}(V) = \begin{bmatrix} A_{\text{K}}(1) & \beta_n(V) & 0 & 0 & 0 \\ 4\alpha_n(V) & A_{\text{K}}(2) & 2\beta_n(V) & 0 & 0 \\ 0 & 3\alpha_n(V) & A_{\text{K}}(3) & 3\beta_n(V) & 0 \\ 0 & 0 & 2\alpha_n(V) & A_{\text{K}}(4) & 4\beta_n(V) \\ 0 & 0 & 0 & \alpha_n(V) & A_{\text{K}}(5) \end{bmatrix}, \quad (5.11)$$

and the diagonal elements

$$A_{\text{ion}}(i) = - \sum_{j: j \neq i} A_{\text{ion}}(j, i), \text{ for ion} \in \{\text{Na}, \text{K}\}.$$

5.3 Relation Between the 14D and 4D Deterministic HH Models

Earnshaw et al [30] suggested that it is reasonable to expect that the global flow of the 14D system should converge to the 4D submanifold but also that it is far from obvious that it must. Existing theory applies to the voltage-clamped case [30, 63]. Here, we consider instead the current-clamped case, in which the fluctuations of the ion channel

state influences the voltage evolution, and *vice-versa*.

In the remainder of this section we will (1) define a *multinomial submanifold* \mathcal{M} and show that it is an invariant manifold within the 14D space, and (2) show that the velocity on the 14D space and the pushforward of the velocity on the 4D space are identical. In §5.4 we will (3) provide numerical evidence that \mathcal{M} is globally attracting within the higher-dimensional space.

In order to compare the trajectories of the 14D HH equations with trajectories of the standard 4D equations, we define lower-dimensional and higher-dimensional domains \mathcal{X} and \mathcal{Y} , respectively, as

$$\begin{aligned}\mathcal{X} &= \{-\infty < v < \infty, 0 \leq m \leq 1, 0 \leq h \leq 1, 0 \leq n \leq 1\} = \mathbb{R} \times [0, 1]^3 \subset \mathbb{R}^4 \\ \mathcal{Y} &= \{-\infty < v < \infty\} \cap \left\{ 0 \leq m_{ij}, \sum_{i=0}^3 \sum_{j=0}^1 m_{ij} = 1 \right\} \cap \left\{ 0 \leq n_i, \sum_{i=0}^4 n_i = 1 \right\} \\ &= \mathbb{R} \times \Delta^7 \times \Delta^4 \subset \mathbb{R}^{14},\end{aligned}\tag{5.12}$$

where Δ^k is the k -dimensional simplex in \mathbb{R}^{k+1} given by $y_1 + \dots + y_{k+1} = 1, y_i \geq 0$. The 4D HH model $\frac{dx}{dt} = F(x)$, equations (5.1)-(5.4), is defined for $x \in \mathcal{X}$, and the 14D HH model $\frac{dy}{dt} = G(y)$, equations (5.7)-(5.9), is defined for $y \in \mathcal{Y}$. We introduce a dimension-

14D model	4D model
$(v, m_{00}, \dots, m_{31}, n_0, \dots, n_4)$	(v, m, h, n)
v	v
$\frac{1}{3}(m_{11} + m_{10}) + \frac{2}{3}(m_{21} + m_{20}) + m_{31} + m_{30}$	m
$m_{01} + m_{11} + m_{21} + m_{31}$	h
$n_1/4 + n_2/2 + 3n_3/4 + n_4$	n

Table 5.1: R : Map from the 14D HH model $(m_{00}, \dots, m_{31}, n_0, \dots, n_4)$ to the 4D HH model (m, h, n) . Note that $\{m_{00}, \dots, m_{31}\}$ and $\{n_0, \dots, n_4\}$ both follow multinomial distributions.

reducing mapping $R : \mathcal{Y} \rightarrow \mathcal{X}$ as in Table 5.1, and a mapping from lower to higher dimension, $H : \mathcal{X} \rightarrow \mathcal{Y}$ as in Table 5.2. We construct R and H in such a way that $R \circ H$ acts as the identity on \mathcal{X} , that is, for all $x \in \mathcal{X}$, $x = R(H(x))$. The maps H and R are consistent with a multinomial structure for the ion channel state distribution, in the following sense. The space \mathcal{Y} covers all possible probability distributions on the eight sodium channel states and the five potassium channel states. Those distributions which are products of one multinomial distribution on the K^+ -channel¹ and a second multinomial distribution on the Na^+ -channel² form a submanifold of $\Delta^7 \times \Delta^4$. In this way we define a submanifold, denoted $\mathcal{M} = H(\mathcal{X})$, the image of \mathcal{X} under H .

4D model	14D model
(v, m, h, n)	$(v, m_{00}, \dots, m_{31}, n_0, \dots, n_4)$
v	v
$(1 - n)^4$	n_0
$4(1 - n)^3 n$	n_1
$6(1 - n)^2 n^2$	n_2
$4(1 - n) n^3$	n_3
n^4	n_4
$(1 - m)^3(1 - h)$	m_{00}
$3(1 - m)^2 m(1 - h)$	m_{10}
$3(1 - m) m^2(1 - h)$	m_{20}
$m^3(1 - h)$	m_{30}
$(1 - m)^3 h$	m_{01}
$3(1 - m)^2 m h$	m_{11}
$3(1 - m) m^2 h$	m_{21}
$m^3 h$	m_{31}

Table 5.2: H : Map from the 4D HH model (m, h, n) and the 14D HH model $(m_{00}, \dots, m_{31}, n_0, \dots, n_4)$.

Before showing that the multinomial submanifold \mathcal{M} is an invariant manifold within

¹That is, distributions indexed by a single open probability n ; with the five states having probabilities $\binom{4}{i} n^i (1 - n)^{4-i}$ for $0 \leq i \leq 4$.

²That is, distributions indexed by two open probabilities m and h , with the eight states having probabilities $\binom{3}{i} m^i (1 - m)^{3-i} h^j (1 - h)^{1-j}$, for $0 \leq i \leq 3$, and $0 \leq j \leq 1$.

the 14D space, we first show that the deterministic 14D HH model is defined on a bounded domain. Having a bounded forward-invariant manifold is a general property of conductance-based models, which may be written in the form

$$\frac{dV}{dt} = f(V, \mathcal{N}_{\text{open}}) = \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_{\text{leak}}) - \sum_{i \in \mathcal{I}} [g_i N_{\text{open}}^i (V - V_i)] \right\} \quad (5.13)$$

$$\frac{d\mathcal{N}}{dt} = A(V)\mathcal{N} \text{ and} \quad (5.14)$$

$$\mathcal{N}_{\text{open}} = \mathcal{O}[\mathcal{N}]. \quad (5.15)$$

Here, C is the membrane capacitance, I_{app} is an applied current with upper and lower bounds I_{\pm} respectively, and g_i is the conductance for the i th ion channel. The index i runs over the set of distinct ion channel types in the model, \mathcal{I} . The gating vector \mathcal{N} represents the fractions of each ion channel population in various ion channel states, and the operator \mathcal{O} gives the fraction of each ion channel population in the open (or conducting) channel states. The following lemma establishes that any conductance-based model (including the 4D or 14D HH model) is defined on a bounded domain.

Lemma 1. *For a conductance-based model of the form (5.13)-(5.15), and for any bounded applied current $I_- \leq I_{\text{app}} \leq I_+$, there exist upper and lower bounds V_{max} and V_{min} such that trajectories with initial voltage condition $V \in [V_{\text{min}}, V_{\text{max}}]$ remain within this interval for all times $t > 0$, regardless of the initial channel state.*

Proof. Let $V_1 = \min_{i \in \mathcal{I}} \{V_i\} \wedge V_{\text{leak}}$, and $V_2 = \max_{i \in \mathcal{I}} \{V_i\} \vee V_{\text{leak}}$, where the index i runs over \mathcal{I} , the set of distinct ion channel types. Note that for all i , $0 \leq N_{\text{open}}^i \leq 1$, and

$g_i > 0$, $g_{\text{leak}} > 0$. Therefore when $V \leq V_1$

$$\frac{dV}{dt} = \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_{\text{leak}}) - \sum_{i \in \mathcal{I}} [g_i N_{\text{open}}^i (V - V_i)] \right\} \quad (5.16)$$

$$\geq \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_1) - \sum_{i \in \mathcal{I}} [g_i N_{\text{open}}^i (V - V_1)] \right\} \quad (5.17)$$

$$\geq \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_1) - \sum_{i \in \mathcal{I}} [g_i \times 0 \times (V - V_1)] \right\} \quad (5.18)$$

$$= \frac{1}{C} \{ I_{\text{app}} - g_{\text{leak}}(V - V_1) \}. \quad (5.19)$$

Inequality (5.17) follows because $V_1 = \min_{i \in \mathcal{I}} \{V_i\} \wedge V_{\text{leak}}$, and inequality (5.18) follows because $V - V_1 \leq 0$, $g_i > 0$ and $N_{\text{open}}^i \geq 0$. Let $V_{\min} := \min \left\{ \frac{I_{\text{app}}}{g_{\text{leak}}} + V_1, V_1 \right\}$. When $V < V_{\min}$, $\frac{dV}{dt} > 0$. Therefore, V will not decrease beyond V_{\min} .

Similarly, when $V \geq V_2$

$$\frac{dV}{dt} = \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_{\text{leak}}) - \sum_{i \in \mathcal{I}} [g_i N_{\text{open}}^i (V - V_i)] \right\} \quad (5.20)$$

$$\leq \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_2) - \sum_{i \in \mathcal{I}} [g_i N_{\text{open}}^i (V - V_2)] \right\} \quad (5.21)$$

$$\leq \frac{1}{C} \left\{ I_{\text{app}} - g_{\text{leak}}(V - V_2) - \sum_{i \in \mathcal{I}} [g_i \times 0 \times (V - V_2)] \right\} \quad (5.22)$$

$$= \frac{1}{C} \{ I_{\text{app}} - g_{\text{leak}}(V - V_2) \}. \quad (5.23)$$

Inequality (5.21) holds because $V_2 = \max_{i \in \mathcal{I}} \{V_i\} \vee V_{\text{leak}}$, and inequality (5.22) holds because $V - V_2 \geq 0$, $g_i > 0$ and $N_{\text{open}}^i \geq 0$. Let $V_{\max} = \max \left\{ \frac{I_{\text{app}}}{g_{\text{leak}}} + V_2, V_2 \right\}$. When $V > V_{\max}$, $\frac{dV}{dt} < 0$. Therefore, V will not go beyond V_{\max} .

We conclude that if V takes an initial condition in the interval $[V_{\min}, V_{\max}]$, then $V(t)$ remains within this interval for all $t \geq 0$. \square

Given that $y \in \mathcal{Y}$ has a bounded domain, Lemma 2 follows directly, and establishes that the multinomial submanifold \mathcal{M} is a forward-time-invariant manifold within the 14D space.

Lemma 2. *Let \mathcal{X} and \mathcal{Y} be the lower-dimensional and higher-dimensional Hodgkin-Huxley manifolds given by (5.12), and let F and G be the vector fields on \mathcal{X} and \mathcal{Y} defined by (5.1)-(5.4) and (5.7)-(5.9), respectively. Let $H : \mathcal{X} \rightarrow \mathcal{M} \subset \mathcal{Y}$ and $R : \mathcal{Y} \rightarrow \mathcal{X}$ be the mappings given in Tables 5.2 and 5.1, respectively, and define the multinomial submanifold $\mathcal{M} = H(\mathcal{X})$. Then \mathcal{M} is forward-time-invariant under the flow generated by G . Moreover, the vector field G , when restricted to \mathcal{M} , coincides with the vector field induced by F and the map H . That is, for all $y \in \mathcal{M}$, $G(y) = D_x H(R(y)) \cdot F(R(y))$.*

The main idea of the proof is to show that for every $y \in \mathcal{Y}$, $G(y)$ is contained in the span of the four vectors $\left\{ \frac{\partial H}{\partial x_i}(R(y)) \right\}_{i=1}^4$.

Proof. The map from the 4D HH model to the 14D HH model is given in Tab. 5.2 as $\{H : x \rightarrow y \mid x \in \mathcal{X}, y \in \mathcal{Y}\}$, and the map from the 14D HH model to the 4D HH model is given in Tab. 5.1 as $\{R : y \rightarrow x \mid x \in \mathcal{X}, y \in \mathcal{Y}\}$. The partial derivatives $\frac{\partial H}{\partial x}$ of the map

H are given by

$$\begin{aligned}\frac{dm_{00}}{dm} &= -3(1-m)^2(1-h) \\ \frac{dm_{10}}{dm} &= 3(1-h)(3m^2-4m+1) \\ \frac{dm_{20}}{dm} &= 3(1-h)(2m-3m^2) \\ \frac{dm_{30}}{dm} &= 3(1-h)m^2 \\ \frac{dm_{01}}{dm} &= -3h(1-m)^2 \\ \frac{dm_{11}}{dm} &= 3h(3m^2-4m+1) \\ \frac{dm_{21}}{dm} &= 3h(2m-3m^2) \\ \frac{dm_{31}}{dm} &= 3hm^2\end{aligned}$$

$$\begin{aligned}\frac{dm_{00}}{dh} &= -(1-m)^3 \\ \frac{dm_{10}}{dh} &= -3(1-m)^2m \\ \frac{dm_{20}}{dh} &= -3(1-m)m^2 \\ \frac{dm_{30}}{dh} &= -m^3 \\ \frac{dm_{01}}{dh} &= (1-m)^3 \\ \frac{dm_{11}}{dh} &= 3(1-m)^2m \\ \frac{dm_{21}}{dh} &= 3(1-m)m^2 \\ \frac{dm_{31}}{dh} &= m^3.\end{aligned}$$

We can write $\partial H/\partial x$ in matrix form as:

$$\frac{\partial H}{\partial x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3(1-m)^2(1-h) & -(1-m)^3 & 0 \\ 0 & 3(1-h)(3m^2-4m+1) & -3(1-m)^2m & 0 \\ 0 & 3(1-h)(2m-3m^2) & -3(1-m)m^2 & 0 \\ 0 & 3(1-h)m^2 & -m^3 & 0 \\ 0 & -3h(1-m)^2 & (1-m)^3 & 0 \\ 0 & 3h(3m^2-4m+1) & 3(1-m)^2m & 0 \\ 0 & 3h(2m-3m^2) & 3(1-m)m^2 & 0 \\ 0 & 3hm^2 & m^3 & 0 \\ 0 & 0 & 0 & -4(1-n)^3 \\ 0 & 0 & 0 & 4(1-n)^2(1-4n) \\ 0 & 0 & 0 & 12n(1-n)(1-2n) \\ 0 & 0 & 0 & 4n^2(3-4n) \\ 0 & 0 & 0 & 4n^3 \end{bmatrix}.$$

We write out the vector fields (5.8) and (5.9) component by component:

$$\begin{aligned} \frac{d\mathbf{M}_1}{dt} &= \beta_m \mathbf{M}_2 + \beta_h \mathbf{M}_5 - (3\alpha_m + \alpha_h) \mathbf{M}_1 \\ &= -3(1-m)^2(1-h) [(1-m)\alpha_m - m\beta_m] + (1-m)^3 [h\beta_h - (1-h)\alpha_h] \\ \frac{d\mathbf{M}_2}{dt} &= 3\alpha_m \mathbf{M}_1 + 2\beta_m \mathbf{M}_3 + \beta_h \mathbf{M}_6 - (2\alpha_m + \beta_m + \alpha_h) \mathbf{M}_2 \\ &= 3(1-h)(3m^2-4m+1) [(1-m)\alpha_m - m\beta_m] + 3(1-m)^2m [h\beta_h - (1-h)\alpha_h] \\ \frac{d\mathbf{M}_3}{dt} &= 2\alpha_m \mathbf{M}_2 + 3\beta_m \mathbf{M}_4 + \beta_h \mathbf{M}_7 - (\alpha_m + 2\beta_m + \alpha_h) \mathbf{M}_3, \\ &= 3(1-h)(2m-3m^2) [(1-m)\alpha_m - m\beta_m] + 3(1-m)m^2 [h\beta_h - (1-h)\alpha_h] \end{aligned}$$

$$\begin{aligned}
\frac{d\mathbf{M}_4}{dt} &= \alpha_m \mathbf{M}_3 + \beta_h \mathbf{M}_8 - (3\beta_m + \alpha_h) \mathbf{M}_4, \\
&= 3(1-h)m^2 [(1-m)\alpha_m - m\beta_m] + m^3 [h\beta_h - (1-h)\alpha_h] \\
\frac{d\mathbf{M}_5}{dt} &= \beta_m \mathbf{M}_6 + \alpha_h \mathbf{M}_1 - (3\alpha_m + \beta_h) \mathbf{M}_5, \\
&= -3h(1-m)^2 [(1-m)\alpha_m - m\beta_m] + (1-m)^3 [h\beta_h - (1-h)\alpha_h] \\
\frac{d\mathbf{M}_6}{dt} &= 3\alpha_m \mathbf{M}_5 + 2\beta_m \mathbf{M}_7 + \alpha_h \mathbf{M}_2 - (2\alpha_m + \beta_m + \beta_h) \mathbf{M}_6, \\
&= 3h(3m^2 - 4m + 1) [(1-m)\alpha_m - m\beta_m] - 3(1-m)^2 m [h\beta_h - (1-h)\alpha_h] \\
\frac{d\mathbf{M}_7}{dt} &= 2\alpha_m \mathbf{M}_6 + 3\beta_m \mathbf{M}_8 + \alpha_h \mathbf{M}_3 - (\alpha_m + 2\beta_m + \beta_h) \mathbf{M}_7, \\
&= 3h(2m - 3m^2) [(1-m)\alpha_m - m\beta_m] - 3(1-m)m^2 [h\beta_h - (1-h)\alpha_h] \\
\frac{d\mathbf{M}_8}{dt} &= \alpha_m \mathbf{M}_7 + \alpha_h \mathbf{M}_4 - (3\beta_m + \beta_h) \mathbf{M}_8, \\
&= 3hm^2 [(1-m)\alpha_m - m\beta_m] - m^3 [h\beta_h - (1-h)\alpha_h] \\
\frac{d\mathbf{N}_1}{dt} &= \beta_n \mathbf{N}_2 - 4\alpha_n \mathbf{N}_1 = -4(1-n)^3 [\alpha_n(1-n) - n\beta_n], \\
\frac{d\mathbf{N}_2}{dt} &= 4\alpha_n \mathbf{N}_1 + 2\beta_n \mathbf{N}_3 - (3\alpha_n + \beta_n) \mathbf{N}_2 = 4(1-n)^2 (1-4n) [\alpha_n(1-n) - n\beta_n], \\
\frac{d\mathbf{N}_3}{dt} &= 3\alpha_n \mathbf{N}_2 + 3\beta_n \mathbf{N}_4 - (2\alpha_n + 2\beta_n) \mathbf{N}_3 = 12n(1-n)(1-2n) [\alpha_n(1-n) - n\beta_n], \\
\frac{d\mathbf{N}_4}{dt} &= 2\alpha_n \mathbf{N}_3 + 4\beta_n \mathbf{N}_5 - (3\alpha_n + 3\beta_n) \mathbf{N}_4 = 4n^2(3-4n) [\alpha_n(1-n) - n\beta_n], \\
\frac{d\mathbf{N}_5}{dt} &= \alpha_n \mathbf{N}_4 - 4\beta_n \mathbf{N}_5 = 4n^3 [\alpha_n(1-n) - n\beta_n].
\end{aligned}$$

By extracting common factors from the previous expressions it is clear that $G(y)$ may be

written, for all $y \in \mathcal{Y}$, as

$$\begin{aligned}
G(y) = & \frac{-\bar{g}_{Na}\mathbf{M}_8(V - V_{Na}) - \bar{g}_K\mathbf{N}_5(V - V_K) - g_L(V - V_L) + I_{\text{app}}}{C} \left\{ \frac{\partial H}{\partial v}(R(y)) \right\} \\
& + [(1 - m')\alpha_m - m'\beta_m] \left\{ \frac{\partial H}{\partial m}(R(y)) \right\} \\
& - [h'\beta_h - (1 - h')\alpha_h] \left\{ \frac{\partial H}{\partial h}(R(y)) \right\} \\
& + [\alpha_n(1 - n') - n'\beta_n] \left\{ \frac{\partial H}{\partial n}(R(y)) \right\}
\end{aligned} \tag{5.24}$$

where $m' = (\mathbf{M}_2 + \mathbf{M}_6)/3 + 2(\mathbf{M}_3 + \mathbf{M}_7)/3 + (\mathbf{M}_4 + \mathbf{M}_8)$, $h' = \mathbf{M}_5 + \mathbf{M}_6 + \mathbf{M}_7 + \mathbf{M}_8$ and $n' = \mathbf{N}_2/4 + \mathbf{N}_3/2 + 3\mathbf{N}_4/4 + \mathbf{N}_5$. Thus $G(y)$ is in the span of the column vectors $\partial H/\partial v$, $\partial H/\partial m$, $\partial H/\partial n$, and $\partial H/\partial h$, as was to be shown.

On the other hand, the vector field for the 4D HH ODE (5.1-5.4) is given by

$$F = \begin{bmatrix} (-\bar{g}_{Na}m^3h(V - V_{Na}) - \bar{g}_Kn^4(V - V_K) - g_L(V - V_L) + I_{\text{app}}) / C \\ \alpha_m(V)(1 - m) - \beta_m(V)m \\ \alpha_h(V)(1 - h) - \beta_h(V)h \\ \alpha_n(V)(1 - n) - \beta_n(V)n \end{bmatrix}.$$

Referring to (5.24), we see that $G(y) = D_x H(R(y)) F(R(y))$. Thus we complete the proof of Lemma 2. \square

Lemma 2 establishes that the 14D HH model given by (5.7)-(5.9) is dynamically consistent with the original 4D HH model given by (5.1)-(5.4). In §5.4 we provide numerical evidence that the flow induced by G on \mathcal{Y} converges to \mathcal{M} exponentially fast. Thus, an initial probability distribution over the ion channel states that is not multinomial quickly approaches a multinomial distribution with dynamics induced by the 4D HH equations. Similar results, restricted to the voltage-clamp setting, were established by Keener and

Earnshaw [30, 63, 65].

5.4 Local Convergence Rate

Keener and Earnshaw [30, 63, 65] showed that for Markov chains with constant (even time varying) transition rates: (i) the multinomial probability distributions corresponding to mean-field models (such as the HH sodium or potassium models) form invariant submanifolds within the space of probability distributions over the channel states, and (ii) arbitrary initial probability distributions converged exponentially quickly to the invariant manifold. For systems with prescribed time-varying transition rates, such as for an ion channel system under voltage clamp with a prescribed voltage $V(t)$ as a function of time, the distribution of channel states had an invariant submanifold again corresponding to the multinomial distributions, and the flow on that manifold induced by the evolution equations was consistent with the flow of the full system.

In the preceding section we established the dynamical consistency of the 14D and 4D models with enough generality to cover both the voltage-clamp and current-clamp systems; the latter is distinguished by NOT having a prescribed voltage trace, but rather having the voltage coevolve along with the (randomly fluctuating) ion channel states. Here, we give numerical evidence for exponential convergence under current clamp similar to that established under voltage clamp by Keener and Earnshaw.

Rather than providing a rigorous proof, we give numerical evidence for the standard deterministic HH model that $y \rightarrow \mathcal{M}$ under current clamp (spontaneous firing conditions) in the following sense: if $y(t)$ is a solution of $\dot{y} = G(y)$ with arbitrary initial data $y_0 \in \mathcal{Y}$, then $\|y(t) - H(R(y(t)))\| \rightarrow 0$ as $t \rightarrow \infty$, exponentially quickly. Moreover, the convergence rate is bounded by $\lambda = \max(\lambda_v, \lambda_{Na}, \lambda_K)$, where λ_{ion} is the least negative

nontrivial eigenvalue of the channel state transition matrix (over the voltage range $V_{\min} \leq v \leq V_{\max}$) for a given ion, and $-1/\lambda_v$ is the largest value taken by the membrane time constant (for $V_{\min} \leq v \leq V_{\max}$). In practice, we find that the membrane time constant does not determine the slowest time scale for convergence to \mathcal{M} . In fact it appears that the second-least-negative eigenvalues (not the least-negative eigenvalues) of the ion channel matrices set the convergence rate.

Note that $y \in \mathcal{Y}$ can be written as $y = [V; \mathbf{M}; \mathbf{N}]$. The Jacobian matrix $\frac{\partial H}{\partial x}$ consists of three block matrices: one for the voltage terms, $\frac{\partial V}{\partial v}$, one associated to the Na^+ gates, given by $\frac{\partial \mathbf{M}}{\partial m}$ and $\frac{\partial \mathbf{M}}{\partial h}$, and one corresponding to the K^+ gates, $\frac{\partial \mathbf{N}}{\partial n}$. Fixing a particular voltage v , let λ_i , $i \in \{0, 1, 2, \dots, 7\}$ be the eight eigenvalues of A_{Na} and v_i be the associated eigenvectors, i.e., $A_{\text{Na}}v_i = \lambda_i v_i$ for the rate matrix in equations (5.8). Similarly, let η_i , w_i , $i \in \{0, 1, 2, \dots, 4\}$ be the five eigenvalues and the associated eigenvectors of A_{K} , i.e., $A_{\text{K}}w_i = \eta_i w_i$, for the rate matrix in equations (5.9). If we rank the eigenvalues of either matrix in descending order, the leading eigenvalue is always zero (because the sum of each column for A_{Na} and A_{K} is zero for every V) and the remainder are real and negative. Let λ_1 and η_1 denote the largest (least negative) nontrivial eigenvalues of A_{Na} and A_{K} , respectively, and let $v_1 \in \mathbb{R}^8$ and $w_1 \in \mathbb{R}^5$ be the corresponding eigenvectors.

The eigenvectors of the full 14D Jacobian are not simply related to the eigenvectors of the component submatrices, because the first (voltage) row and column contain nonzero off-diagonal elements. However, the eigenvectors associated to the largest nonzero eigenvalues of A_{Na} and A_{K} (respectively v_2 and w_2) are parallel to $\partial M/\partial h$ and $\partial N/\partial n$, regardless of voltage. In other words, the slowest decaying directions for each ion channel, v_1 and w_1 , transport the flow along the multinomial sub-manifold of \mathcal{Y} . Therefore, it is reasonable to make the hypothesis that if $Y(t)$ is a solution of $\dot{y} = G(y)$ with arbitrary initial data $y \in \mathcal{Y}$,

then

$$\frac{\|y(t) - H(R(y(t)))\|}{\|y(0) - H(R(y(0)))\|} \lesssim e^{-\lambda_2 t} \quad (5.25)$$

for λ_2 being the *second largest* nonzero eigenvalue of A_K and A_{Na} over all v in the range $v_{\min} < v < v_{\max}$. The convergence behavior is plotted numerically in Fig. 5.2, and is consistent with the Ansatz (5.25). We calculate the distance from a point y to \mathcal{M} as

$$y_{\max} = \operatorname{argmax}_{y \in \mathcal{Y}} \|y - H(R(y))\|^2. \quad (5.26)$$

In order to obtain an upper bound on the distance as a function of time, we begin with the furthest point in the simplex from \mathcal{M} , by numerically finding the solution to the argument (5.26), which is

$$y_{\max} = [v, 0.5, 0, 0, 0.5, 0, 0, 0, 0, 0.5, 0, 0, 0, 0.5].$$

This vector represents the furthest possible departure from the multinomial distribution: all probability equally divided between the extreme states m_{00} and m_{03} for the sodium channel, and the extremal states n_0 and n_4 for potassium. The maximum distance from the multinomial submanifold \mathcal{M} , d_{\max} , is calculated using this point. As shown in Fig. 5.2, the function $d_{\max} e^{-\lambda_2 t}$ provides a tight upper bound for the convergence rate from arbitrary initial data $y \in \mathcal{Y}$ to the invariant submanifold \mathcal{M} .

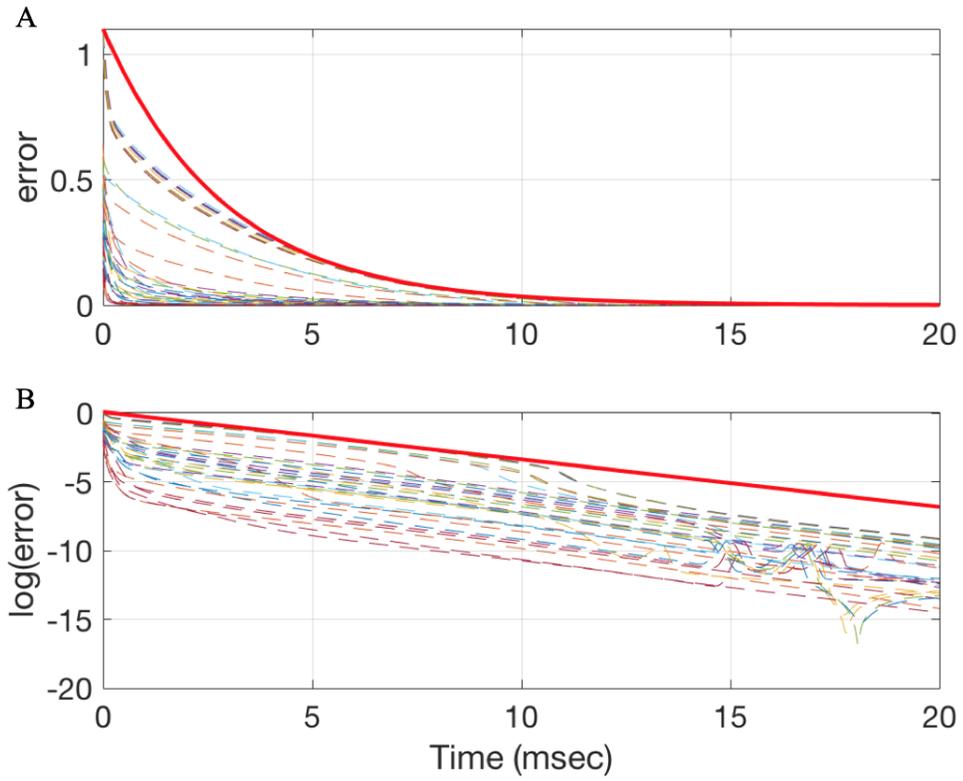


Figure 5.2: Convergence of trajectories $y(t)$, for arbitrary initial conditions $y_0 \in \mathcal{Y}$, to the multinomial submanifold \mathcal{M} , for an ensemble of random initial conditions. A: distance (eqn. (5.26)) between $y(t)$ and \mathcal{M} . B: Logarithm of the distance in panel A. The red solid line shows $d_{\max} e^{-\lambda_2 t}$ in panel A and $\log(d_{\max}) - \lambda_2 t$ in panel B.

Chapter 6

Stochastic 14D Hodgkin-Huxley Models

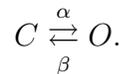
Finite populations of ion channels generate stochastic fluctuations (“channel noise”) in ionic currents that influence action potential initiation and timing [107, 125]. At the molecular level, fluctuations arise because transitions between individual ion channel states are stochastic [58, 85, 116]. Each directed edge in the ion channel state transition diagrams (cf. Fig. 3.1) introduces an independent noise source. It is of interest to be able to attribute variability of the interspike interval timing distribution to specific molecular noise sources, specifically individual directed edges in each channel state graph. In order to explore these contributions, we develop a system of Langevin equations for the Hodgkin-Huxley equations, set in a 14-dimensional phase space.

Working with a higher-dimensional stochastic model may appear inconvenient, but in fact has several advantages. First, any projection of an underlying 14D model onto a lower (e.g. 4D) stochastic model generally entails loss of the Markov property. Second, the higher-dimensional representation allows us to assess the contribution of individual molecular transitions to the macroscopically observable variability of timing in the interspike

interval distribution. Third, by using a rectangular noise coefficient matrix constructed directly from the transitions in the ion channel graphs, we avoid a matrix decomposition step.¹ This approach leads to a fast algorithm that is equivalent to the slower algorithm due to Fox and Lu [39, 49] in a strong sense (pathwise equivalence) that we detail in §7.

6.1 Two Stochastic Representations for Stochastic Ion Channel Kinetics

Before introducing Langevin models for the stochastic HH kinetics, we will review two stochastic representations for stochastic ion channel kinetics [4]. First, we consider a system with two possible states, C (“closed”) and O (“open”) and assume that the dwell times in C and O are determined by independent exponential random variables with parameters $\alpha > 0$ and $\beta > 0$, as shown below:



Given an initial state $x_0 \in \{C, O\}$, denote $p_{x_0}(x, t)$ to be the probability of being in state $x \in \{C, O\}$ at time t . The “chemical master equation” for this system is given by

$$\frac{d}{dt}p_{x_0}(C, t) = -\alpha p_{x_0}(C, t) + \beta p_{x_0}(O, t) \quad (6.1)$$

$$\frac{d}{dt}p_{x_0}(O, t) = \alpha p_{x_0}(C, t) - \beta p_{x_0}(O, t). \quad (6.2)$$

The chemical master equation provides two linear ordinary differential equations that govern the dynamical behavior of the probability distribution of the model, which is not a

¹In [49], they used the singular value decomposition (SVD) to calculate the square root of the diffusion matrix and Fox [41] proposed to use a Cholesky decomposition. We will discuss the numerical behavior of both in the Discussion (cf. §12).

stochastic representation for the underlying process.

To construct a pathwise representation, [4, 5] constructed a counting process as follows. Let $R_1(t)$ be the number of times the transition $C \rightarrow O$ has taken place by time t and $R_2(t)$ be the number of times the transition $O \rightarrow C$ has taken place by time t . Moreover, let $X_1(t) \in \{0, 1\}$ equal one if the channel is closed at time t , and zero if open; and $X_2(t) \in \{0, 1\}$ take the value one if the channel is open at time t , and zero if close. Denoting $X(t) = (X_1(t), X_2(t))^T$, then

$$X(t) = X(0) + R_1(t) \begin{pmatrix} -1 \\ 1 \end{pmatrix} + R_2(t) \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

To represent the counting process for $X(t)$, [5] constructed a unit-rate Poisson process as follows. Let $\{e_i\}_{i=1}^{\infty}$ be independent exponential random variables with a parameter of one and $\tau_1 = e_1, \tau_2 = \tau_1 + e_2, \dots, \tau_n = \tau_{n-1} + e_n, \dots$. The associated unit-rate Poisson process, $Y(s)$, is the counting process determined by the number of points $\{\tau_i\}_{i=1}^{\infty}$ before $s \geq 0$. For instance, if we let "x" denote the points τ_n in the image below,

x	\dots						
					s		

then $Y(s)=5$. The random time representation defines a non-constant rate $\lambda(s) : [0, \infty) \rightarrow \mathbb{R}_{\geq 0}$, and the number of points observed by time s is $Y \left(\int_0^s \lambda(r) dr \right)$. When $\lambda(s) > 0$, the probability of seeing a jump within the next small increment of time Δs is

$$P \left(Y \left(\int_0^{s+\Delta s} \lambda(r) dr \right) - Y \left(\int_0^s \lambda(r) dr \right) \geq 1 \right) = \lambda(s)\Delta s + o(\Delta s). \quad (6.3)$$

Therefore, the propensity of seeing the next jump is $\lambda(s)$. For the two-state model, the

propensities of the open reaction ($C \rightarrow O$) and close reaction ($O \rightarrow C$) are

$$\lambda_1(X(s)) = \alpha X_1(s), \quad \lambda_2(X(s)) = \beta X_2(s).$$

These facts imply that

$$R_1(t) = Y_1\left(\int_0^t \lambda_1(X(s))ds\right) = Y_1\left(\int_0^t \alpha X_1(s)ds\right),$$

$$R_2(t) = Y_2\left(\int_0^t \lambda_2(X(s))ds\right) = Y_2\left(\int_0^t \beta X_2(s)ds\right).$$

Therefore, a pathwise representation for the stochastic two-state model is

$$X(t) = X(0) + Y_1\left(\int_0^t \alpha X_1(s)ds\right) \begin{pmatrix} -1 \\ 1 \end{pmatrix} + Y_2\left(\int_0^t \beta X_2(s)ds\right) \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (6.4)$$

where Y_1 and Y_2 are two independent, unit-rate Poisson processes. This is the *random time change representation* for this particular reaction system.

The transition rates in the two-state example we discussed here are constant, but in general they can be time dependent or state dependent. For example, Markov chain models of ion channels typically have voltage-dependent propensities. For HH-like ion channel kinetics, the propensity of the channel-opening reaction is

$$\lambda_1(t, X(t)) = \alpha(t)X_1(t),$$

and the propensity of the channel-closing reaction is

$$\lambda_2(t, X(t)) = \beta(t)X_2(t),$$

where $\alpha(t)$, and $\beta(t)$ are voltage-dependent non-negative functions. The analogous expression for the pathwise representation [4] can be expressed as

$$X(t) = X(0) + Y_1 \left(\int_0^t \alpha(s) X_1(s) ds \right) \begin{pmatrix} -1 \\ 1 \end{pmatrix} + Y_2 \left(\int_0^t \beta(s) X_2(s) ds \right) \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (6.5)$$

As a generalization, [4] extended the random time change representation to a d -state model with M reaction channels. Denote $X_i(t)$ to be the value of the i th component at time t , with $X(t) \in \mathbb{Z}^d$ and the propensity of the k th reaction being $\lambda_k(t, X(t))$. If we further assume that the system is updated immediately once the k th reaction happens at time t according to the reaction vector $\zeta_k \in \mathbb{Z}^d$, then

$$X(t) = X(t-) + \zeta_k.$$

Thus, the expression for the pathwise representation is

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(s, X(s)) ds \right) \zeta_k, \quad (6.6)$$

where the Y_k are independent unit-rate Poisson processes.

Algorithm for the simulation of the random time change representation The time until the next reaction occurs after time T is given by

$$\Delta = \min_k \left\{ \Delta_k : \int_0^{T+\Delta_k} \lambda_k(s, X(s)) ds = \tau_T^k \right\},$$

where τ_T^k is the first point associated with Y_k coming after $\int_0^T \lambda_k(s, X(s))ds$, which is defined as follows

$$\tau_T^k \equiv \inf \left\{ r > \int_0^T \lambda_k(s, X(s))ds : Y_k(r) - Y_k \left(\int_0^T \lambda_k(s, X(s))ds \right) = 1 \right\}.$$

For a single ion channel type with state variable X , and including the membrane potential V as an additional dynamical variable, then its pathwise representation is given by

$$C \frac{dV}{dt} = I_{app}(t) - I_V(V(t)) - \left(\sum_{i=1}^d g_i^o X_i(t) \right) (V(t) - V_X). \quad (6.7)$$

Here g_i^o is the conductance of an individual channel when it is in the i th state (of channel X) and the sum gives the total conductance associated with the channel represented by the vector X , and V_X is the reversal potential for the ion channel. Meanwhile the channel state $X(t)$ evolves stochastically according to (6.6).

Algorithm 1. *Simulation random time change representation*

1. *Initialize: set the initial number of molecules of each species, X . Set the initial voltage value V . Set $t = 0$. For each k , set $\tau_k = 0$ and $T_k = 0$.*
2. *Generate M independent, uniform(0,1) random numbers $\{r_k\}_{k=1}^M$, for each $k \in \{1, 2, \dots, M\}$, set $\tau_k = \ln(1/r_k)$*
3. *Numerically integrate equation (6.7) forward in time until one of the following equalities hold:*

$$\int_t^{t+\Delta} \lambda_k(V(s), X(s))ds = \tau_k - T_k. \quad (6.8)$$

4. *Let μ be the index of the reaction channel where the equality (6.8) holds.*

5. For each k , set

$$T_k = T_k + \int_t^{t+\Delta} \lambda_k(V(s), X(s)) ds,$$

where Δ is determined in step 3.

6. Set $t = t + \Delta$ and $X = X + \zeta_\mu$

7. Let r be uniform(0,1) and set $\tau_\mu = \tau_\mu + \ln(1/r)$.

8. Return to step 3 or stop.

Gillespie's algorithm A more widely used algorithm for simulating the stochastic ion channel kinetics is the Gillespie's algorithm [16, 46], which is a time-homogeneous method. For simulating the same system for a multi-state single ion channel kinetics in equation (6.7), the Gillespie's algorithm can be presented as follows. Let Y be a unit rate Poisson process, and $\{\xi_i, i = 0, 1, 2, \dots\}$ be independent, uniform (0, 1) random variables that are independent of Y . If we define

$$\lambda_0(V(s), X(s)) = \sum_{k=1}^M \lambda_k(V(s), X(s)), \quad (6.9)$$

$$q_k(s) = \frac{1}{\lambda_0(V(s), X(s))} \sum_{l=1}^k \lambda_l(V(s), X(s)), \quad (6.10)$$

where $q_0 = 0$ and $k \in \{1, 2, \dots, M\}$. The representation for the ‘‘Gillespie algorithm’’ for the stochastic process of X and V can be written as

$$R_0(t) = Y \left(\int_0^t \lambda_0(V(s), X(s)) ds \right) \quad (6.11)$$

$$X(t) = X(0) + \sum_{k=1}^M \xi_k \int_0^t \mathbb{1}\{\xi_{R_0(s-)} \in [q_{k-1}(s-), q_k(s-)]\} dR_0(s) \quad (6.12)$$

$$C \frac{dV}{dt} = I_{app}(t) - I_V(V(t)) - \left(\sum_{i=1}^d g_i^o X_i(t) \right) (V(t) - V_X). \quad (6.13)$$

The corresponding algorithm is given by

Algorithm 2. *The Gillespie’s algorithm*

1. *Initialize: set the initial number of molecules of each species, X . Set a initial voltage value V and $t = 0$.*
2. *Let r be uniform(0,1) and numerically integrate equation (6.11) forward in time until:*

$$\int_t^{t+\Delta} \lambda_0(V(s), X(s)) ds = \ln(1/r).$$
3. *Let ξ be uniform(0,1) and find the index k where $\xi \in [q_{k-1}((t+\Delta)-), q_k((t+\Delta)-)]$*
4. *Set $t = t + \Delta$ and $X = X + \zeta_k$*
5. *Return to step 2 or stop*

6.2 Exact Stochastic Simulation of HH Kinetics: the Random–Time-Change Representation

An ‘‘exact’’ representation of the Hodgkin-Huxley system with a population of M_{tot} sodium channels and N_{tot} potassium channels treats each of the 20 directed edges in the sodium

channel diagram, and each of the 8 directed edges in the potassium channel diagram, as independent Poisson processes, with voltage-dependent per capita intensities. As in the deterministic case, the sodium and potassium channel population vectors \mathbf{M} and \mathbf{N} satisfy $\sum_{i=0}^3 \sum_{j=0}^1 \mathbf{M}_{ij} \equiv 1 \equiv \sum_{i=0}^4 \mathbf{N}_i$.² Thus they are constrained, respectively, to a 7D simplex embedded in \mathbb{R}^8 and a 4D simplex embedded in \mathbb{R}^5 . In the random–time-change representation [6] (which we follow closely throughout this section) the exact evolution equations are written in terms of sums over the directed edges \mathcal{E} for each ion channel, $\mathcal{E}_{\text{Na}} = \{1, \dots, 20\}$ and $\mathcal{E}_{\text{K}} = \{1, \dots, 8\}$, cf. Fig. 3.1.

$$\mathbf{M}(t) = \mathbf{M}(0) + \frac{1}{M_{\text{tot}}} \sum_{k \in \mathcal{E}_{\text{Na}}} \zeta_k^{\text{Na}} Y_k^{\text{Na}} \left(M_{\text{tot}} \int_0^t \alpha_k^{\text{Na}}(V(s)) \mathbf{M}_{i(k)}(s) ds \right) \quad (6.14)$$

$$\mathbf{N}(t) = \mathbf{N}(0) + \frac{1}{N_{\text{tot}}} \sum_{k \in \mathcal{E}_{\text{K}}} \zeta_k^{\text{K}} Y_k^{\text{K}} \left(N_{\text{tot}} \int_0^t \alpha_k^{\text{K}}(V(s)) \mathbf{N}_{i(k)}(s) ds \right). \quad (6.15)$$

Here ζ_k^{ion} is the stoichiometry vector for the k th directed edge. If we write $i(k)$ for the source node and $j(k)$ for the destination node of edge k , then $\zeta_k^{\text{ion}} = e_{j(k)}^{\text{ion}} - e_{i(k)}^{\text{ion}}$.³ Each $Y_k^{\text{ion}}(\tau)$ is an independent unit-rate Poisson process, evaluated at “internal time” (or integrated intensity) τ , representing the independent channel noise arising from transitions along the k th edge. The voltage-dependent per capita transition rate along the k th edge is $\alpha_k^{\text{ion}}(v)$, and $\mathbf{M}_{i(k)}(s)$ (resp. $\mathbf{N}_{i(k)}(s)$) is the fractional occupancy of the source node for the k th transition at time s . Thus, for example, the quantity $M_{\text{tot}} \alpha_k^{\text{Na}}(V(s)) \mathbf{M}_{i(k)}(s)$ gives the *net* intensity along the k th directed edge in the Na^+ channel graph at time s .

Remark 1. Under “voltage-clamp” conditions, with the voltage V held fixed, (6.14)-(6.15) reduce to a time-invariant first-order transition process on a directed graph [42, 105].

²We annotate the stochastic population vector \mathbf{M} either as $[M_{00}, M_{10}, \dots, M_{31}]$ or as $[M_1, \dots, M_8]$, whichever is more convenient. In either notation $M_{31} \equiv M_8$ is the conducting state of the Na^+ channel. For the K^+ channel, N_4 denotes the conducting state.

³We write e_i^{Na} and e_i^{K} for the i th standard unit vector in \mathbb{R}^8 or \mathbb{R}^5 , respectively.

Under “current-clamp” conditions, the voltage evolves according to a conditionally deterministic current balance equation of the form

$$\frac{dV}{dt} = \frac{1}{C} \{I_{\text{app}}(t) - \bar{g}_{\text{Na}}\mathbf{M}_{31}(V - V_{\text{Na}}) - \bar{g}_{\text{K}}\mathbf{N}_4(V - V_{\text{K}}) - g_{\text{leak}}(V - V_{\text{leak}})\}. \quad (6.16)$$

Here, C ($\mu\text{F}/\text{cm}^2$) is the capacitance, I_{app} ($n\text{A}/\text{cm}^2$) is the applied current, the maximal conductance is \bar{g}_{chan} ($m\text{S}/\text{cm}^2$), V_{chan} ($m\text{V}$) is the associated reversal potential, and the ohmic leak current is $g_{\text{leak}}(V - V_{\text{leak}})$.

The random–time-change representation (6.14)-(6.16) leads to an exact stochastic simulation algorithm, given in [6]; equivalent simulation algorithms have been used previously [18, 87]. Many authors substitute a simplified Gillespie algorithm that imposes a piecewise-constant propensity approximation, described above, that ignores the voltage dependence of the transition rates α_k^{ion} between channel transition events [48, 49, 90, 93]. The two methods give similar moment statistics, provided $N_{\text{tot}}, M_{\text{tot}} \gtrsim 40$ [6]; their similarity regarding path-dependent properties (including interspike interval distributions) has not been studied in detail. Moreover, both Markov chain algorithms are prohibitively slow for modest numbers (e.g. thousands) of channels; the exact algorithm may be even slower than the approximate Gillespie algorithm. For consistency with previous studies, in this thesis we use the piecewise-constant propensity Gillespie algorithm with $M_{\text{tot}} = 6000 \text{ Na}^+$ and $N_{\text{tot}} = 1800 \text{ K}^+$ channels as our “gold standard” Markov chain (MC) model, as in [49].

In §6.3 we develop a 14D conductance-based Langevin model with 28 independent noise sources – one for each directed edge – derived from the random–time-change representation (6.14)-(6.16). In previous work [105] we established a quantitative measure of “edge importance”, namely the contribution of individual transitions (directed edges) to the variance of channel state occupancy under steady-state voltage-clamp conditions.

Under voltage clamp, the edge importance was identical for each reciprocal pair of directed edges in the graph, a consequence of detailed balance. Some Langevin models lump the noise contributions of each pair of edges [20, 21, 90, 93]. Under conditions of detailed balance, this simplification is well justified. However, as we will show (cf. Fig. 6.2) under current-clamp conditions, e.g. for an actively spiking neuron, detailed balance is violated, the reciprocal edge symmetry is broken, and each pair of directed edges makes a distinct contribution to ISI variability.

6.3 Previous Langevin Models

As reviewed by [49, 61, 73], the general Langevin models based on Fox and Lu's work [39] (cf. eqn (3.1)-(3.3)) for HH kinetics incorporates noise into the deterministic HH equations as follows:

$$C\dot{V} = -g_K(n^4 + \phi_K)(V - V_K) - g_{Na}(m^3h + \phi_{Na})(V - V_{Na}) - g_L(V - V_L) + I, \quad (6.17)$$

$$\dot{n} = \alpha_n(1 - n) - \beta_n n + \eta_n, \quad (6.18)$$

$$\dot{m} = \alpha_m(1 - n) - \beta_m m + \eta_m, \quad (6.19)$$

$$\dot{h} = \alpha_h(1 - h) - \beta_h h + \eta_h, \quad (6.20)$$

where ϕ_K , ϕ_{Na} , η_n , η_m and η_h are mean zero noise terms, the exact form of which depends on the model we consider.

Fox and Lu's model (1997)

In this model [40], $\phi_K^F \equiv \phi_{Na}^F \equiv 0$, where the superscript F is used to indicate the case of this model (same as the following models). Terms η_n , η_m and η_h are Gaussian white noise with mean squares

$$\langle \eta_n(t), \eta_n(t') \rangle^F = \frac{\alpha_n(1-n) + \beta_n n}{N_K} \delta(t-t'), \quad (6.21)$$

$$\langle \eta_m(t), \eta_m(t') \rangle^F = \frac{\alpha_m(1-m) + \beta_m m}{N_{Na}} \delta(t-t'), \quad (6.22)$$

$$\langle \eta_h(t), \eta_h(t') \rangle^F = \frac{\alpha_h(1-h) + \beta_h h}{N_{Na}} \delta(t-t'), \quad (6.23)$$

where N_K and N_{Na} denotes the number of potassium and sodium channels.

Linaro et al. model (2011)

In the paper [73], they proposed a Langevin model where

$$\phi_K^L = \sum_{i=1}^4 z_{K,i}, \quad (6.24)$$

and

$$\tau_{K,i} \dot{z}_{K,i} = -z_{K,i} + \sigma_{K,i} \sqrt{2\tau_{K,i}} \xi_{K,i}. \quad (6.25)$$

Here, $\tau_{K,i}$ and $\sigma_{K,i}$ are some functions of the opening and closing rates of the n-gates. The random processes $\xi_{K,i}$ represent independent delta-correlated Gaussian white noise with zero mean and unitary variances. Similarly, the noise term for Na^+ channels reads

$$\phi_{Na}^L = \sum_{i=1}^7 z_{Na,i}, \quad (6.26)$$

where

$$\tau_{\text{Na},i} \dot{z}_{\text{Na},i} = -z_{\text{Na},i} + \sigma_{\text{Na},i} \sqrt{2\tau_{\text{Na},i}} \xi_{\text{Na},i}. \quad (6.27)$$

Moreover, the noise term for gating variables are zero: $\eta_n \equiv \eta_m \equiv \eta_h \equiv 0$.

Güler and Marifi's model (2013)

In [56], the authors mentioned the “nontrivial cross-correlation persistence (NCCP)”, which they claimed to be the major cause of the elevated excitability and spontaneous firing in small-size neuronal membranes. The term ϕ_{K}^G , which reflects NCCP attributed to the K^+ channels, is given by

$$\phi_{\text{K}}^G = \sqrt{\frac{n^4(1-n^4)}{N_{\text{K}}}} q_{\text{K}}, \quad (6.28)$$

where

$$\tau \dot{q}_{\text{K}} = p_{\text{K}}, \quad (6.29)$$

$$\tau \dot{p}_{\text{K}} = -\gamma_{\text{K}} p_{\text{K}} - \omega_{\text{K}}^2 [\alpha_n(1-n) + \beta_n n] q_{\text{K}} + \xi_{\text{K}}, \quad (6.30)$$

and ξ_{K} is a delta-correlated, mean zero Gaussian white noise term with mean squares

$$\langle \xi_{\text{K}}(t), \xi_{\text{K}}(t') \rangle^G = \gamma_{\text{K}} T_{\text{K}} [\alpha_n(1-n) + \beta_n n] \delta(t-t'). \quad (6.31)$$

In the above equations, τ corresponds to the unit time. The constants are dimensionless with: $\gamma_{\text{K}} = 10$, $\omega_{\text{K}}^2 = 150$ and $T_{\text{K}} = 400$. Similarly, the term ϕ_{Na}^G , which captures NCCP attributed to the Na^+ channels reads

$$\phi_{\text{Na}}^G = \sqrt{\frac{m^3(1-m^3)}{N_{\text{Na}}}} h q_{\text{Na}}, \quad (6.32)$$

where

$$\tau \dot{q}_{\text{Na}} = p_{\text{Na}}, \quad (6.33)$$

$$\tau \dot{p}_{\text{Na}} = -\gamma_{\text{Na}} p_{\text{Na}} - \omega_{\text{Na}}^2 [\alpha_m(1-m) + \beta_m m] q_{\text{Na}} + \xi_{\text{Na}}, \quad (6.34)$$

with ξ_{Na} a mean zero Gaussian white noise term with mean squares

$$\langle \xi_{\text{Na}}(t), \xi_{\text{Na}}(t') \rangle^G = \gamma_{\text{Na}} T_{\text{Na}} [\alpha_m(1-m) + \beta_m m] \delta(t-t'). \quad (6.35)$$

In the above equations, τ corresponds to the unit time. The constants are dimensionless with: $\gamma_{\text{Na}} = 10$, $\omega_{\text{Na}}^2 = 200$ and $T_{\text{Na}} = 800$.

The subunit noises are Gaussian and their mean squares satisfy

$$\langle \eta_n(t), \eta_n(t') \rangle^F = \frac{\alpha_n(1-n) + \beta_n n}{4N_{\text{K}}} \delta(t-t'), \quad (6.36)$$

$$\langle \eta_m(t), \eta_m(t') \rangle^F = \frac{\alpha_m(1-m) + \beta_m m}{3N_{\text{Na}}} \delta(t-t'), \quad (6.37)$$

$$\langle \eta_h(t), \eta_h(t') \rangle^F = \frac{\alpha_h(1-h) + \beta_h h}{N_{\text{Na}}} \delta(t-t'), \quad (6.38)$$

where N_{K} and N_{Na} denote the number of potassium and sodium channels.

Orio and Soudry's Model (2012)

In [90], Orio and Soudry derived the SDE explicitly for HH type ion channel kinetic. They examined a specific population of N ion channels with M states, where the transition rate of a single channel from state j to i is given by A_{ji} . In addition, they define the matrix A to be composed of A_{ij} terms for all $i \neq j$, and $A_{ii} = -\sum_{j \neq i} A_{ji}$ on the diagonal. They denote by x_i the fraction of channels in each of the state, and by \mathbf{x} a vector of x_i . The

diffusion approximation (DA) for the stochastic dynamics of \mathbf{x} leads to the same form of SDE as Fox and Lu's 94 model

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x} + S\xi, \quad (6.39)$$

where ξ is a vector of independent Gaussian white noise processes with zero mean and unit variance, A is the rate matrix (same as eqn (5.10)-(5.11)). Goldwyn et al [49] provided a numerical method for Fox and Lu's 94 model with a great computational cost, of order $O(M^3)$ at each time step. In [90], Orío and Soudry proposed a more efficient method, where

$$D_{ij} = \frac{1}{N} \begin{cases} \sum_{k \neq i} (A_{ik}x_k(t) + A_{ki}x_i(t)), & \text{if } i = j \\ -A_{ji}x_i(t) - A_{ij}x_j(t), & \text{if } i \neq j \end{cases}$$

They denote by T the set of all possible transitions pairs ($i \rightleftharpoons j$) that exist between states and then give each pair an index in $k = 1, \dots, |T|$. Note that $|T|$ is the size of set T . Also, they denote $T(i)$ to be the subset of all transitions pairs that connect to state i , and m_{ik} to be the index of the state connected by the k^{th} transition pair, excluding state i . Finally, they write matrix S of size $M \times |T|$, and

$$S_{ik} = \begin{cases} \text{sign}(i - m_{ik}) \frac{1}{\sqrt{N}} \sqrt{A_{(i,m_{ik})}x_{m_{ik}} + A_{(m_{ki},i)}x_i} & , k \in T(i) \\ 0 & , k \notin T(i). \end{cases}$$

Note that Dangerfield *et al.* [20] proposed a Langevin model that shares the same construction as the Orío and Soudry's model. In [20, 21], they implemented the model with different boundary conditions than Orío and Soudry's model, which we will discuss in detail in §8.

6.4 The $14 \times 28\text{D}$ HH Model

For sufficiently large number of channels, [104, 105] showed that under voltage clamp, equations (6.14)-(6.15) can be approximated by a multidimensional Ornstein-Uhlenbeck (OU) process (or Langevin equation) in the form⁴

$$d\mathbf{M} = \sum_{k=1}^{20} \zeta_k^{\text{Na}} \left\{ \alpha_k^{\text{Na}}(V) \mathbf{M}_{i(k)} dt + \sqrt{\epsilon^{\text{Na}} \alpha_k^{\text{Na}}(V) \mathbf{M}_{i(k)}} dW_k^{\text{Na}} \right\} \quad (6.40)$$

$$d\mathbf{N} = \sum_{k=1}^8 \zeta_k^{\text{K}} \left\{ \alpha_k^{\text{K}}(V) \mathbf{N}_{i(k)} dt + \sqrt{\epsilon^{\text{K}} \alpha_k^{\text{K}}(V) \mathbf{N}_{i(k)}} dW_k^{\text{K}} \right\}. \quad (6.41)$$

Here, \mathbf{M} , \mathbf{N} , ζ_k^{ion} , and α_k^{ion} have the same meaning as in (6.14)-(6.15). The channel state *increments* in a short time interval dt are $d\mathbf{M}$ and $d\mathbf{N}$, respectively. The finite-time increment in the Poisson process Y_k^{ion} is now approximated by a Gaussian process, namely the increment dW_k^{ion} in a Wiener (Brownian motion) process associated with each directed edge. These independent noise terms are scaled by $\epsilon^{\text{Na}} = 1/M_{\text{tot}}$ and $\epsilon^{\text{K}} = 1/N_{\text{tot}}$, respectively.

Equations (6.16)-(6.41) comprise a system of Langevin equations for the HH system (under current clamp) on a 14-dimensional phase space driven by 28 independent white noise sources, one for each directed edge. These equations may be written succinctly in the form

$$d\mathbf{X} = \mathbf{f}(\mathbf{X}) dt + \sqrt{\epsilon} \mathcal{G}(\mathbf{X}) d\mathbf{W}(t) \quad (6.42)$$

where we define the 14-component vector $\mathbf{X} = (V; \mathbf{M}; \mathbf{N})$, and $\mathbf{W}(t)$ is a Wiener process with 28 independent components. The deterministic part of the evolution equation $\mathbf{f}(\mathbf{X}) = \left[\frac{dV}{dt}; \frac{d\mathbf{M}}{dt}; \frac{d\mathbf{N}}{dt} \right]$ is the same as the mean-field, equations (5.7)-(5.9). The state-dependent

⁴The convergence of the discrete channel system to a Langevin system under voltage clamp is a special case of Kurtz' theorem [71].

noise coefficient matrix \mathcal{G} is 14×28 and can be written as

$$\sqrt{\epsilon}\mathcal{G} = \left(\begin{array}{c|c} \mathbf{0}_{1 \times 20} & \mathbf{0}_{1 \times 8} \\ \hline S_{\text{Na}} & \mathbf{0}_{8 \times 8} \\ \hline \mathbf{0}_{5 \times 20} & S_{\text{K}} \end{array} \right).$$

The matrices S_{K} and S_{Na} are given by

$$S_{\text{K}} = \left[\begin{array}{cccc} -\sqrt{4\alpha_n n_0} & \sqrt{\beta_n n_1} & 0 & 0 \\ \sqrt{4\alpha_n n_0} & -\sqrt{\beta_n n_1} & -\sqrt{3\alpha_n n_1} & \sqrt{2\beta_n n_2} \\ 0 & 0 & \sqrt{3\alpha_n n_1} & -\sqrt{2\beta_n n_2} \cdots \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ \cdots & & & \\ \sqrt{2\alpha_n n_2} & -\sqrt{3\beta_n n_3} & -\sqrt{\alpha_n n_3} & \sqrt{4\beta_n n_4} \\ 0 & 0 & \sqrt{\alpha_n n_3} & -\sqrt{4\beta_n n_4} \end{array} \right],$$

and

$$\begin{aligned}
 S_{\text{Na}}^{(1:5)} &= \begin{bmatrix} -\sqrt{\alpha_h m_{00}} & \sqrt{\beta_h m_{01}} & -\sqrt{3\alpha_m m_{00}} & \sqrt{\beta_m m_{10}} & 0 \\ 0 & 0 & \sqrt{3\alpha_m m_{00}} & -\sqrt{\beta_m m_{10}} & -\sqrt{\alpha_h m_{10}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \sqrt{\alpha_h m_{00}} & -\sqrt{\beta_h m_{01}} & 0 & 0 & 0 \\ -\sqrt{\beta_h m_{11}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\alpha_h m_{20}} & -\sqrt{\beta_h m_{21}} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
 S_{\text{Na}}^{(6:10)} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \sqrt{\beta_h m_{11}} & -\sqrt{2\alpha_m m_{10}} & \sqrt{2\beta_m m_{20}} & 0 & 0 \\ \sqrt{2\alpha_m m_{10}} & -\sqrt{2\beta_m m_{20}} & -\sqrt{\alpha_h m_{20}} & \sqrt{\beta_h m_{21}} & \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\sqrt{\beta_h m_{11}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\alpha_h m_{20}} & -\sqrt{\beta_h m_{21}} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

$$S_{\text{Na}}^{(11:15)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\sqrt{\alpha_m m_{20}} & \sqrt{3\beta_m m_{30}} & 0 & 0 & 0 \\ \sqrt{\alpha_m m_{20}} & -\sqrt{3\beta_m m_{30}} & -\sqrt{\alpha_h m_{30}} & \sqrt{\beta_h m_{31}} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{3\alpha_m m_{01}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\alpha_h m_{30}} & -\sqrt{\beta_h m_{31}} & 0 \end{bmatrix},$$

$$S_{\text{Na}}^{(16:20)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\sqrt{3\alpha_m m_{01}} & \sqrt{\beta_m m_{11}} & 0 & 0 & 0 \\ -\sqrt{\beta_m m_{11}} & -\sqrt{2\alpha_m m_{11}} & \sqrt{2\beta_m m_{21}} & 0 & 0 \\ 0 & \sqrt{2\alpha_m m_{11}} & -\sqrt{2\beta_m m_{21}} & -\sqrt{\alpha_m m_{21}} & \sqrt{3\beta_m m_{31}} \\ 0 & 0 & 0 & \sqrt{\alpha_m m_{21}} & -\sqrt{3\beta_m m_{31}} \end{bmatrix},$$

where $S_{\text{Na}}^{(i:j)}$ is the i^{th} - j^{th} column of S_{Na} .

Note that each of the 8 columns of S_K corresponds to the flux vector along a single directed edge in the K^+ channel transition graph. Similarly, each of the 20 columns of S_{Na} corresponds to the flux vector along a directed edge in the Na^+ graph (cf. Fig. 3.1).

Remark 2. *Although the ion channel state trajectories generated by equation (6.42) are not strictly bounded to remain within the nonnegative simplex, empirically, the voltage nevertheless remains within specified limits with overwhelming probability.*

To facilitate comparison of the model (6.16)-(6.41) with prior work [39, 49, 40], we may rewrite the 14×28 D Langevin description in the equivalent form

$$C \frac{dV}{dt} = I_{\text{app}}(t) - \bar{g}_{\text{Na}} \mathbf{M}_8 (V - V_{\text{Na}}) - \bar{g}_{\text{K}} \mathbf{N}_5 (V - V_{\text{K}}) - g_{\text{leak}}(V - V_{\text{leak}}), \quad (6.43)$$

$$\frac{d\mathbf{M}}{dt} = A_{\text{Na}} \mathbf{M} + S_{\text{Na}} \xi_{\text{Na}}, \quad (6.44)$$

$$\frac{d\mathbf{N}}{dt} = A_{\text{K}} \mathbf{M} + S_{\text{K}} \xi_{\text{K}}, \quad (6.45)$$

The drift matrices A_{Na} and A_{K} remain the same as in [39], and are the same as in the 14D deterministic model (5.10)-(5.11). S_{Na} and S_{K} are constructed from direct transitions of the underlying kinetics in Fig. 3.1, $\xi_{\text{Na}} \in \mathbb{R}^{20}$ and $\xi_{\text{K}} \in \mathbb{R}^8$ are vectors of independent Gaussian white noise processes with zero mean and unit variance.

Fox and Lu's original approach [39] requires solving a matrix square root equation $SS^\top = D$ to obtain a square (8×8 for Na^+ or 5×5 for K^+) noise coefficient matrix consistent with the state-dependent diffusion matrix D . As an advantage, the ion channel representation (6.43)-(6.45) uses sparse, nonsquare noise coefficient matrices (8×20 for the Na^+ channel and 5×8 for the K^+ channel), which exposes the independent sources of noise for the system.

The new Langevin model in (6.43)-(6.45) does not require detailed balance, which gives more insights to the underlying kinetics. Review papers such as [49, 61, 93], did systematic comparison of various stochastic versions of the HH model. In §7 and §8, we quantitatively analyze the connection between the new model and other existing models [20, 21, 39, 40, 41, 48, 49, 60, 61, 90, 93]. Problems such as the boundary constraints are beyond the scope of this dissertation, however, we would like to connect the new model to another type of approximation to the MC model, namely the stochastic shielding approximation.

6.5 Stochastic Shielding for the 14D HH Model

The stochastic shielding (SS) approximation was introduced by Schmandt and Galán [102], in order to approximate the Markov process using fluctuations from only a subset of the transitions, namely those corresponding to changes in the observable states. In [105], we showed that, under voltage clamp, each directed edge makes a distinct contribution to the steady-state variance of the ion channel conductance, with the total variance being a sum of these contributions. We call the variance due to the k th directed edge the *edge importance*; assuming detailed balance, it is given by

$$R_k = J_k \sum_{i=2}^n \sum_{j=2}^n \left(\frac{-1}{\lambda_i + \lambda_j} \right) (\mathbf{c}^\top v_i) (w_i^\top \zeta_k) (\zeta_k^\top w_j) (v_j^\top \mathbf{c}). \quad (6.46)$$

Here, J_k is the steady-state probability flux along the k th directed edge; $\lambda_n < \lambda_{n-1} \leq \dots \leq \lambda_2 < 0$ are the eigenvalues of the drift matrix (A_{Na} or A_{K} , respectively), and v_i (resp. w_i) are the corresponding right (resp. left) eigenvectors of the drift matrix. Each edge's stoichiometry vector ζ_k has components summing to zero; consequently the columns of A_{Na} and A_{K} all sum to zero. Thus each drift matrix has a leading trivial eigenvalue $\lambda_1 \equiv 0$. The vector \mathbf{c} specifies the unitary conductance of each ion channel state; for the HH model it is proportional to e_8^{Na} or e_5^{K} , respectively.

Fig. 6.1 shows the edge importance for each pair of edges in the HH Na^+ and K^+ ion channel graph, as a function of voltage in the range $[-100, 100]$ mV. Note that reciprocal edges have identical R_k due to detailed balance. Under voltage clamp, the largest value of R_k for the HH channels always corresponds to directly observable transitions, i.e. edges k such that $|\mathbf{c}^\top \zeta_k| > 0$, although this condition need not hold in general [104].

To apply the stochastic shielding method under current clamp, we simulate the model with noise from only a selected subset $\mathcal{E}' \subset \mathcal{E}$ of directed edges, replacing (6.44)-(6.45)

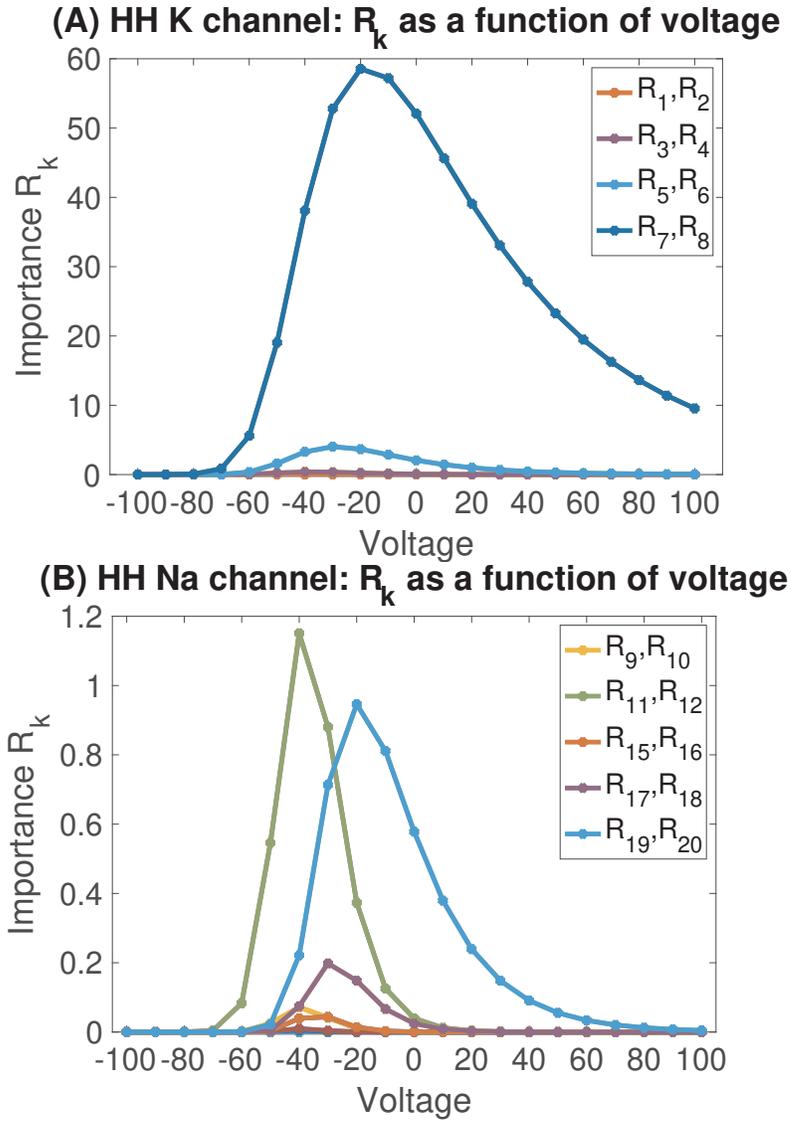


Figure 6.1: Stochastic shielding under voltage clamp. Redrawn (with permission) from Figs. 10 & 13 of [105]. Each curve shows the edge importance R_k (equation (6.46)) as a function of voltage in the range $[-100, 100]$ mV for a different edge pair. For the K^+ kinetics, $R_7 = R_8$ are the largest R_k value in the voltage range above. For the Na^+ kinetics, $R_{11} = R_{12}$ have the largest R_k values in the subthreshold voltage range (c. $[-100, -25]$ mV), and $R_{19} = R_{20}$ have the largest R_k values in the suprathreshold voltage range (c. $[-25, 100]$ mV).

with

$$\frac{d\mathbf{M}}{dt} = A_{\text{Na}}\mathbf{M} + S'_{\text{Na}}\xi_{\text{Na}}, \quad (6.47)$$

$$\frac{d\mathbf{N}}{dt} = A_{\text{K}}\mathbf{M} + S'_{\text{K}}\xi_{\text{K}}, \quad (6.48)$$

where S'_{Na} (resp. S'_{K}) is a reduced matrix containing only the noise coefficients from the most important edges \mathcal{E}' . That is, \mathcal{E}' contains a subset of edges with the largest edge-importance values R_k .

In [102], they assumed that the edges with the largest contribution contribution to current fluctuations under voltage clamp would also make the largest contributions to variability in voltage and timing under current clamp, and included edges 7 – 8 of the K^+ channel ($\mathcal{E}'_{\text{K}} = \{7, 8\}$) and edges 11 – 12 and 19 – 20 of the Na^+ channel ($\mathcal{E}'_{\text{Na}} = \{11, 12, 19, 20\}$), yielding an 8×4 matrix S'_{Na} and an 5×2 matrix S'_{K} . They demonstrated numerically that restricting stochastic forcing to these edges gave a significantly faster simulation with little appreciable change in statistical behavior: under voltage clamp, the mean current remained the same, with a small (but noticeable) decrease in the current variance; meanwhile similar inter-spike interval (ISI) statistics were observed.

Under current clamp, detailed balance is violated, and it is not clear from mathematical principles whether the edges with the largest R_k under voltage clamp necessarily make the largest contribution under other circumstances. In order to evaluate the contribution of the fluctuations driven by each directed edge on ISI variability, we test the stochastic shielding method by removing all but one column of S'_{ion} at a time. That is, we restrict to a single noise source and observe the resulting ISI variance empirically. For example, to calculate the importance of the k^{th} direct edge in the Na^+ channel, we suppress the noise from all

other edges by setting $S'_{\mathcal{K}\mathcal{K}} = \mathbf{0}_{5 \times 1}$ and

$$S'_{\text{Na}} = \left[\mathbf{0}_{8 \times 1}, \dots, S_{\text{Na}}(:, k), \dots, \mathbf{0}_{8 \times 1} \right]$$

i.e., only include the k^{th} column of S_{Na} and set other columns to be zeros. The ISI variance was calculated from an ensemble of 10^4 voltage traces, each spanning c. 500 ISIs.

Fig. 6.2A plots the logarithm of the ISI variance for each edge in $\mathcal{E}_{\mathcal{K}}$. Vertical bars (cyan) show the ensemble mean of the ISI variance, with a 95% confidence interval superimposed (magenta). Several observations are in order. First, the ISI variance driven by the noise in each edge decreases rapidly, the further the edge is from the observable transitions (edges 7,8), reflecting the underlying “stochastic shielding” phenomenon. Second, the symmetry of the edge importance for reciprocal edge pairs ((1,2), (3,4), (5,6) and (7,8)) that is observed under voltage clamp is broken under current clamp. The contribution of individual directed edges to timing variability under current clamp has another important difference compared with the edge importance (current fluctuations) under voltage clamp. A similar breaking of symmetry for reciprocal edges is seen for the Na^+ channel, again reflecting the lack of detailed balance during active spiking.

Fig. 6.2B shows the ISI variance when channel noise is included on individual edges of \mathcal{E}_{Na} . Here the difference between voltage and current clamp is striking. Under voltage clamp, the four most important edges are always those representing “observable transitions”, in the sense that the transition’s stoichiometry vector ζ is not orthogonal to the conductance vector \mathbf{c} . That is, the four most important pairs are always 11-12 and 19-20, regardless of voltage (Fig. 6.1). Under current clamp, the most important edges are 17, 18, 19 and 20. Although edges 11 and 12 are among the four most important sources of channel population fluctuations under voltage clamp, they are not even among the top ten

contributors to ISI variance, when taken singly. Even though edges 17 and 18 are “hidden”, meaning they do not directly change the instantaneous channel conductance, these edges are nevertheless the second most important pair under current clamp. Therefore, when we implement the stochastic-shielding based approximation, we include the pairs 17-18 and 19-20 in equation (6.47). We refer to the approximate SS model driven by these six most important edges as the 14×6 D HH model.

Given the other parameters we use for the HH model (cf. Tab. B.1 in Appendix B), the input current of $I_{\text{app}} = 10$ nA is slightly beyond the region of multistability associated with a subcritical Andronov-Hopf bifurcation. In order to make sure the results are robust against increases in the applied current, we tried current injections ranging from 20 to 100 nA. While injecting larger currents decreased the ISI variance, it did not change the rank order of the contributions from the most important edges (cf. Fig. 6.3).

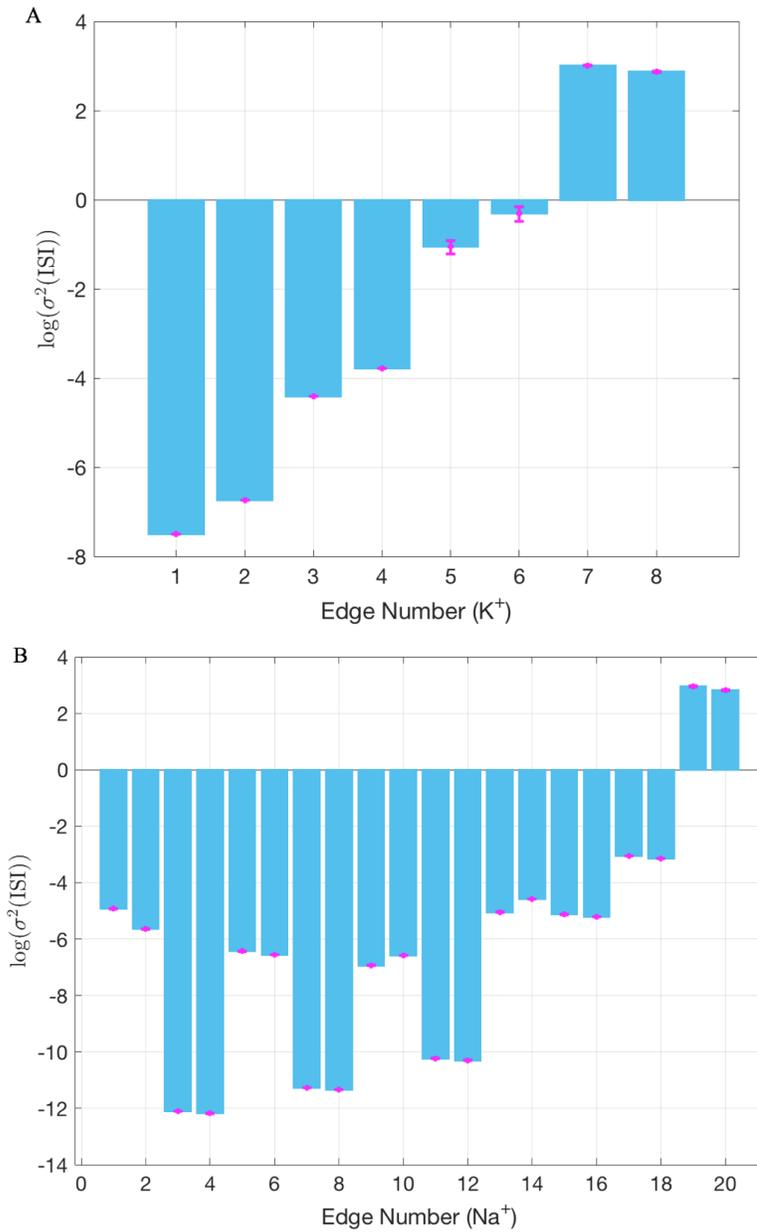


Figure 6.2: Logarithm of variance of ISI for stochastic shielding under current clamp. Cyan bar is the mean of ISI, and magenta plots 95% confidence interval of the mean ISI (see text for details). The applied current is 10 nA with other parameters specified in the Appendices. For the K^+ kinetics, the largest contribution edge is 7, and 8 is slightly smaller ranking the second largest. For the Na^+ kinetics, the largest contribution pair is 19 and 20, with 20 slightly smaller than 19. Moreover, edge 17 and 18 is the second largest pair.

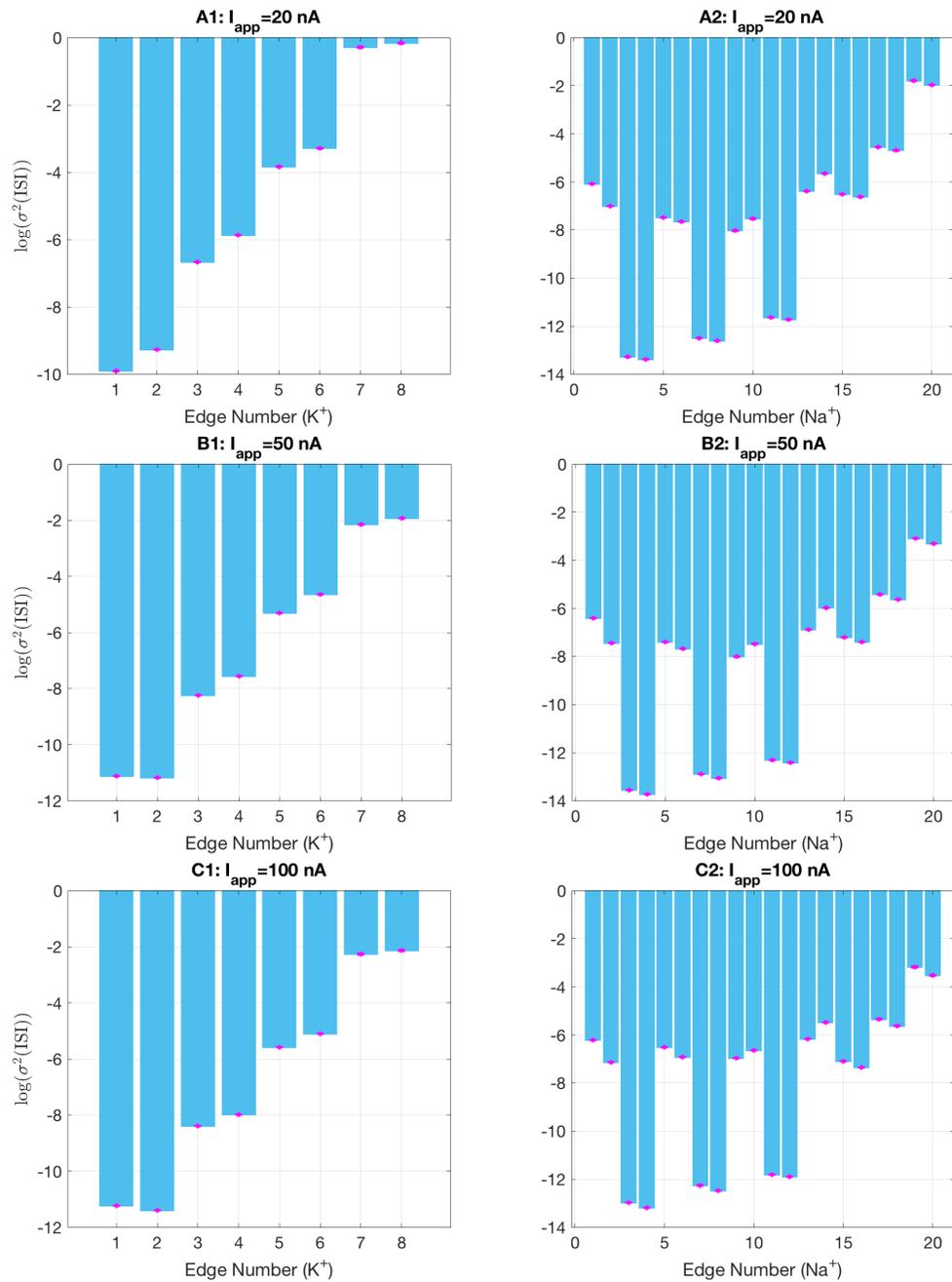


Figure 6.3: Robustness of edge-importance ordering under current clamp for K^+ (column **A1-C1**) and Na^+ (column **A2-C2**). Under each set of applied currents, the largest contribution to ISI variance comes from edge 7 followed by edge 8, of the K^+ kinetics. For the Na^+ kinetics, the largest contribution comes from pair 19-20, and edges 17 and 18 form the second largest pair for Na^+ .

Part III

Model Comparison

Chapter 7

Pathwise Equivalence for a Class of Langevin Models

Fox and Lu’s method was widely used since its appearance (see references in [49, 12, 61]), and the “best” approximation for the underlying Markov Chain (MC) model has been a subject of ongoing discussion for decades. Several studies [12, 81, 111] attested to discrepancies between Fox’s later approach in [40] and the discrete-state MC model, raising the question of whether Langevin approximations could ever accurately represent the underlying fluctuations seen in the “gold standard” MC models. An influential review paper [49] found that these discrepancies were due to the way in which noise is added to the stochastic differential equations (3.1)-(3.3). Recent studies including [20, 21, 41, 48, 49, 56, 60, 61, 73, 90, 93] discussed various ways of incorporating channel noise into HH kinetics based on the original work by Fox and Lu [39, 40], some of which have the same SDEs but with different boundary conditions. Different boundary conditions (BCs) are not expected to have much impact on computational efficiency. Indeed, if BCs are neglected,

the main difference between channel-based (or conductance-based) models is the diffusion matrix S in the Langevin equations (3.2) and (3.3). As the discussion about where and how to incorporate noise into the HH model framework goes on, [41] recently asked whether there is a way of relating different models with different S matrices. We give a positive answer to this question below.

In §7.1 we will demonstrate the equivalence (neglecting the boundary conditions) of a broad class of previously proposed channel-based Langevin models including: [20, 21, 39, 41, 49, 60, 90, 93] and the 14D Langevin HH model with 28 independent noise sources (one for each directed edge in the channel state transition graph), i.e. our “14 × 28D” Langevin model.

7.1 When are Two Langevin Equations Equivalent?

Two Langevin models are pathwise equivalent if the sample paths (trajectories) of one model can be made to be identical to the sample paths of the other, under an appropriate choice of Gaussian white noise samples for each. To make this notion precise, consider two channel-based Langevin models of the form $d\mathbf{X} = \mathbf{f}(\mathbf{X}) dt + G(\mathbf{X}) d\mathbf{W}$ with the same mean dynamics $\mathbf{f} \in \mathbb{R}^d$ and two different $d \times n$ matrices (possibly with different values of n_1 and n_2), G_1 and G_2 . Denote

$$\mathbf{f} : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad (7.1)$$

$$G_1 : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times n_1}, \quad (7.2)$$

$$G_2 : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times n_2}. \quad (7.3)$$

Let $\mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_d(t)]^\top$ and $\mathbf{X}^*(t) = [X_1^*(t), X_2^*(t), \dots, X_d^*(t)]^\top$ be trajectories produced by the two models and let $\mathbf{W}(t) = [W_1(t), W_2(t), \dots, W_{n_1}(t)]^\top$ and $\mathbf{W}^*(t) = [W_1^*(t), W_2^*(t), \dots, W_{n_2}^*(t)]^\top$ be vectors of Wiener processes. That is, $W_i(t)$, $i = 1, 2, \dots, n_1$ and $W_j^*(t)$, $j = 1, 2, \dots, n_2$ are independent Wiener processes with $\langle W_i(s)W_j(t) \rangle = \delta_{ij}\delta(t-s)$ and $\langle W_i^*(s)W_j^*(t) \rangle = \delta_{ij}\delta(t-s)$. Note that n_1 and n_2 need not be equal. As defined in [2], the stochastic differential equation (SDE) models

$$d\mathbf{X} = \mathbf{f}(t, \mathbf{X}(t))dt + G_1(t, \mathbf{X}(t))d\mathbf{W}(t) \quad (7.4)$$

and

$$d\mathbf{X}^* = \mathbf{f}(t, \mathbf{X}^*(t))dt + G_2(t, \mathbf{X}^*(t))d\mathbf{W}^*(t) \quad (7.5)$$

are *pathwise equivalent* if systems (7.4) and (7.5) possess the same probability distribution, and moreover, a sample path solution of one equation is also a sample solution to the other one. In [2], the authors proved a theorem giving general conditions under which the trajectories of two SDEs are equivalent. We follow their construction closely below, adapting it to the case of two different Langevin equations for the Hodgkin-Huxley system represented in a 14-dimensional state space.

As in §6, channel-based Langevin models for the stochastic dynamics of HH can be written as

$$d\mathbf{X} = \mathbf{f}(\mathbf{X}) dt + \mathcal{S}(\mathbf{X}) d\mathbf{W}(t) \quad (7.6)$$

where the 14-component random vector $\mathbf{X} = (V; \mathbf{M}; \mathbf{N})$ and $\mathbf{f}(\mathbf{x}) = \left[\frac{dV}{dt}; \frac{d\mathbf{M}}{dt}; \frac{d\mathbf{N}}{dt} \right]$ is the

same as the mean-field, eqns. (5.7)-(5.9). Recall that $\mathbf{x} = [v, \mathbf{m}, \mathbf{n}]^\top$. Here we write

$$\mathcal{S}(\mathbf{x}) = \left(\begin{array}{c|c} \mathbf{0}_{1 \times m} & \mathbf{0}_{1 \times n} \\ \hline S_{\text{Na}}(\mathbf{m}) & \mathbf{0}_{8 \times n} \\ \hline \mathbf{0}_{5 \times m} & S_{\text{K}}(\mathbf{n}) \end{array} \right), \quad \text{with}$$

$$S_{\text{Na}} : \mathbb{R}^8 \rightarrow \mathbb{R}^{8 \times m}, \quad (7.7)$$

for the Na^+ channel, and

$$S_{\text{K}} : \mathbb{R}^5 \rightarrow \mathbb{R}^{5 \times n}, \quad (7.8)$$

for the K^+ channel. Here, m is the number of independent white noise forcing terms affecting the sodium channel variables, while n is the number of independent noise sources affecting the potassium gating variables. We write

$$\mathbf{W}(t) = [W_1(t), W_2(t), \dots, W_{m+n}(t)]^\top$$

for a Wiener process incorporating both the sodium and potassium noise forcing. Given two channel-based models with diffusion matrices

$$S_{\text{Na},1} : \mathbb{R}^8 \rightarrow \mathbb{R}^{8 \times m_1}, \quad (7.9)$$

$$S_{\text{Na},2} : \mathbb{R}^8 \rightarrow \mathbb{R}^{8 \times m_2}, \quad (7.10)$$

for the Na⁺ channel, and

$$S_{K,1} : \mathbb{R}^5 \rightarrow \mathbb{R}^{5 \times n_1}, \quad (7.11)$$

$$S_{K,2} : \mathbb{R}^5 \rightarrow \mathbb{R}^{5 \times n_2}, \quad (7.12)$$

for the K⁺ channel, we construct the diffusion matrix $\mathcal{D} = \mathcal{S}\mathcal{S}^\top$. In order for the two models to generate equivalent sample paths, it suffices that they have the same diffusion matrix, i.e.

$$\mathcal{D} = \mathcal{S}_1 \mathcal{S}_1^\top = \left(\begin{array}{c|c|c} \mathbf{0}_{1 \times 1} & \mathbf{0}_{1 \times 8} & \mathbf{0}_{1 \times 5} \\ \hline \mathbf{0}_{8 \times 1} & D_{\text{Na}} & \mathbf{0}_{8 \times 5} \\ \hline \mathbf{0}_{5 \times 1} & \mathbf{0}_{5 \times 8} & D_{\text{K}} \end{array} \right) = \mathcal{S}_2 \mathcal{S}_2^\top.$$

The SDEs corresponding to the two channel-based Langevin models are

$$d\mathbf{X} = \mathbf{f}(t, \mathbf{X}(t))dt + \mathcal{S}_1(t, \mathbf{X}(t))d\mathbf{W}(t), \quad (7.13)$$

$$d\mathbf{X}^* = \mathbf{f}(t, \mathbf{X}^*(t))dt + \mathcal{S}_2(t, \mathbf{X}^*(t))d\mathbf{W}^*(t). \quad (7.14)$$

The probability density function $p(t, \mathbf{x})$ for random variable \mathbf{X} in eqn. (7.13) satisfies the Fokker-Planck equation

$$\begin{aligned} \frac{\partial p(t, \mathbf{x})}{\partial t} &= \frac{1}{2} \sum_{i=1}^8 \sum_{j=1}^8 \frac{\partial^2}{\partial x_i \partial x_j} \left[p(t, \mathbf{x}) \sum_{l=1}^{m_1+n_1} \mathcal{S}_1^{(i,l)}(t, \mathbf{x}) \mathcal{S}_1^{(j,l)}(t, \mathbf{x}) \right] \\ &\quad - \sum_{i=1}^8 \frac{\partial}{\partial x_i} \left[\mathbf{f}_i(t, \mathbf{x}) p(t, \mathbf{x}) \right] \\ &= \frac{1}{2} \sum_{i=1}^8 \sum_{j=1}^8 \frac{\partial^2}{\partial x_i \partial x_j} \left[\mathcal{D}^{(i,j)}(t, \mathbf{x}) p(t, \mathbf{x}) \right] \\ &\quad - \sum_{i=1}^8 \frac{\partial}{\partial x_i} \left[\mathbf{f}_i(t, \mathbf{x}) p(t, \mathbf{x}) \right] \end{aligned} \quad (7.15)$$

where $\mathcal{S}_1^{(i,j)}(t, \mathbf{x})$ is the $(i, j)^{th}$ entry of the diffusion matrix $\mathcal{S}_1(t, \mathbf{x})$. Eqn. (7.15) holds because

$$\mathcal{D}^{(i,j)}(t, \mathbf{x}) = \sum_{l=1}^{m_1+n_1} \mathcal{S}_1^{(i,l)}(t, \mathbf{x})\mathcal{S}_1^{(j,l)}(t, \mathbf{x}).$$

If $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^{14}$ and $\mathbf{z}_1 \leq \mathbf{z}_2$, then

$$P(\mathbf{z}_1 \leq \mathbf{X}(t) \leq \mathbf{z}_2) = \int_{z_{1,14}}^{z_{2,14}} \int_{z_{1,13}}^{z_{2,13}} \cdots \int_{z_{1,1}}^{z_{2,1}} p(t, \mathbf{x}) dx_1 dx_2 \cdots dx_8.$$

Note that (7.13) and (7.14) have the same expression (7.15) for the Fokker-Planck equation, therefore, \mathbf{X} and \mathbf{X}^* possess the same probability density function. In other words, the probability density function of \mathbf{X} in eqn. (7.6) is invariant for different choices of the diffusion matrix \mathcal{S} .

7.2 Map Channel-based Langevin Models to Fox and Lu's Model

We now explicitly construct a mapping between Fox and Lu's 14D model [39] and any channel-based model (given the same boundary conditions). We begin with a channel-based Langevin description

$$d\mathbf{X} = \mathbf{f}(t, \mathbf{X}(t))dt + \mathcal{S}(t, \mathbf{X}(t))d\mathbf{W}(t), \quad (7.16)$$

and Fox and Lu's model [39]

$$d\mathbf{X}^* = \mathbf{f}(t, \mathbf{X}^*(t))dt + \mathcal{S}_0(t, \mathbf{X}^*(t))d\mathbf{W}^*(t), \quad (7.17)$$

where S is a d by m matrix satisfying $SS^\top = \mathcal{D}$ (note that S is not necessarily a square matrix), and $\mathcal{S}_0 = \sqrt{\mathcal{D}}$.

Let T be the total simulation time of the random process in equations (7.16) and (7.17). For $0 \leq t \leq T$, denote the singular value decomposition (SVD) of S as

$$S(t) = P(t)\Lambda(t)Q(t)$$

where $P(t)$ is an $d \times d$ orthogonal matrix (i.e., $P^\top P = PP^\top = I_d$) and $Q(t)$ is an $m \times m$ orthogonal matrix, and $\Lambda(t)$ is a $d \times m$ matrix with $\text{rank}(\Lambda) = r \leq d$ positive diagonal entries and $d - r$ zero diagonal entries.

First, we prove that given a Wiener trajectory, $\mathbf{W}(t)$, $t \in [0, T]$ and the solution to eqn. (7.16), $\mathbf{X}(t)$, there exists a Wiener trajectory $\mathbf{W}^*(t)$ such that the solution to eqn. (7.17), \mathbf{X}^* , is also a solution to eqn. (7.16). In other words, for a Wiener process $\mathbf{W}(t)$ we can construct a $\mathbf{W}^*(t)$, such that $\mathbf{X}^*(t) = \mathbf{X}(t)$, for $0 \leq t \leq T$.

Following [2], we construct the vector $\mathbf{W}^*(t)$ of d independent Wiener processes as follows:

$$\mathbf{W}^*(t) = \int_0^t P(s) \left[(\Lambda(s)\Lambda^\top(s))^{\frac{1}{2}} \right]^+ \Lambda(s)Q(s) d\mathbf{W}(s) + \int_0^t P(s) d\mathbf{W}^{**}(s) \quad (7.18)$$

for $0 \leq t \leq T$, where $\mathbf{W}^{**}(t)$ is a vector of length d with the first r entries equal to 0 and the next $d - r$ entries independent Wiener processes, and $\left[(\Lambda(s)\Lambda^\top(s))^{\frac{1}{2}} \right]^+$ is the

pseudoinverse of $(\Lambda(s)\Lambda^\top(s))^{\frac{1}{2}}$. Consider that

$$\mathcal{D}(t) = \mathcal{S}(t)\mathcal{S}^\top(t) = P(t)\Lambda(t)Q(t) \left[P(t)\Lambda(t)Q(t) \right]^\top \quad (7.19)$$

$$= P(t)\Lambda(t)\Lambda^\top(t)P^\top(t) \quad (7.20)$$

$$= [\mathcal{S}_0(t)]^2, \quad (7.21)$$

where $\mathcal{S}_0(t) = P(t) \left(\Lambda(t)\Lambda^\top(t) \right)^{\frac{1}{2}} P^\top(t)$ is a square root of \mathcal{D} , by construction.

The diffusion term on the right side of (7.17) with $\mathbf{X}^*(t)$ replaced by $\mathbf{X}(t)$ satisfies

$$\begin{aligned} & \mathcal{S}_0(t, \mathbf{X}(t))d\mathbf{W}^*(t) \\ &= \mathcal{S}_0(t) \left(P(t) \left[(\Lambda(t)\Lambda^\top(t))^{\frac{1}{2}} \right]^+ \Lambda(t)Q(t)d\mathbf{W}(t) + P(t)d\mathbf{W}^{**}(t) \right) \\ &= P(t) \left(\Lambda(t)\Lambda^\top(t) \right)^{\frac{1}{2}} P^\top(t)P(t) \left[(\Lambda(t)\Lambda^\top(t))^{\frac{1}{2}} \right]^+ \Lambda(t)Q(t)d\mathbf{W}(t) \\ &\quad + P(t) \left(\Lambda(t)\Lambda^\top(t) \right)^{\frac{1}{2}} P^\top(t)P(t)d\mathbf{W}^{**}(t) \\ &= \{P(t)\Lambda(t)Q(t)\} d\mathbf{W}(t). \end{aligned} \quad (7.22)$$

From the SVD of $\mathcal{S}=P\Lambda Q$, we conclude that

$$\mathcal{S}_0(t, \mathbf{X}(t))d\mathbf{W}^*(t) = \mathcal{S}(t, \mathbf{X}(t))d\mathbf{W}(t). \quad (7.23)$$

Hence, $d\mathbf{X} = \mathbf{f}(t, \mathbf{X}(t))dt + \mathcal{S}_0(t, \mathbf{X}(t))d\mathbf{W}_t^*$, i.e., $\mathbf{X}(t)$ is a sample path solution of equation (7.17).

Similarly, given a Wiener trajectory $\mathbf{W}^*(t)$ and the solution to eqn. (7.17) $\mathbf{X}^*(t)$, we can construct a vector $\mathbf{W}(t)$ of m independent Wiener processes as

$$\mathbf{W}(t) = \int_0^t Q^\top(s)\Lambda^+(s) \left[\Lambda(s)\Lambda^\top(s) \right]^{1/2} P^\top(s)d\mathbf{W}^*(s) + \int_0^t Q^\top(s)d\mathbf{W}^{***}(s) \quad (7.24)$$

for $0 \leq t \leq T$, where $\mathbf{W}^{***}(t)$ is a vector of length m with the first r entries equal to 0 and the next $m - r$ entries independent Wiener processes, and $\Lambda^+(s)$ is the pseudoinverse of $\Lambda(s)$. Then, by an argument parallel to (7.22), we conclude that

$$\mathcal{S}(t, \mathbf{X}^*(t))dW(t) = \mathcal{S}_0(t, \mathbf{X}^*(t))d\mathbf{W}^*(t). \quad (7.25)$$

Hence, $d\mathbf{X}^* = \mathbf{f}(t, \mathbf{X}^*(t))dt + \mathcal{S}(t, \mathbf{X}^*(t))d\mathbf{W}^*(t)$, that is, $\mathbf{X}^*(t)$ is also a solution to (7.16). Therefore we can conclude that the channel-based Langevin model in eqn. (7.16) is pathwise equivalent to the Fox and Lu's model.

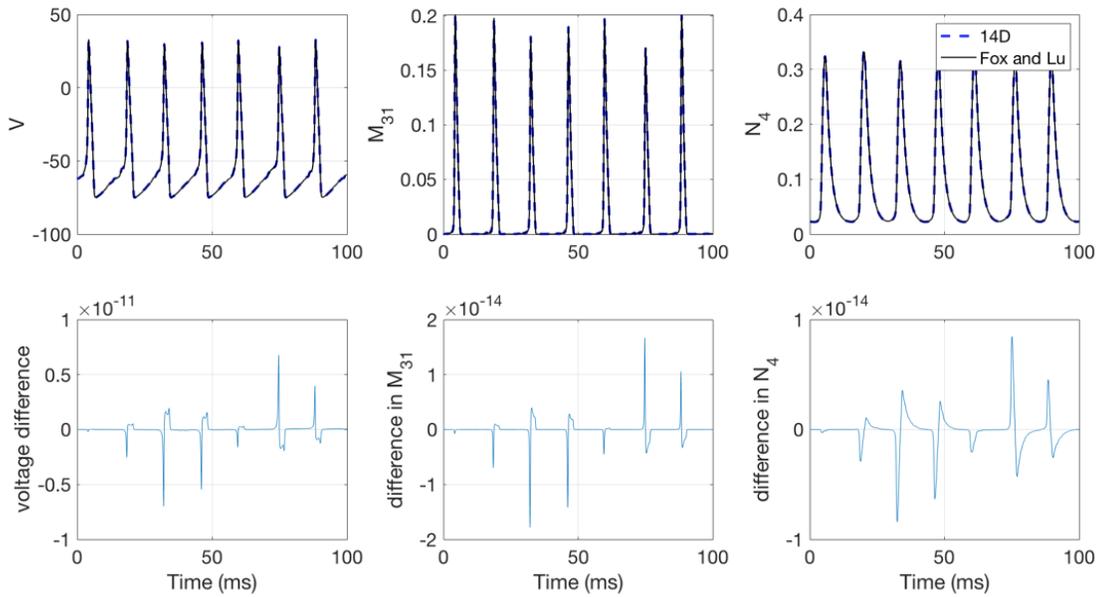
To illustrate pathwise equivalence, Fig. 7.1 plots trajectories of the $14 \times 28\text{D}$ stochastic HH model and Fox and Lu's model, using noise traces dictated by the preceding construction. In panel A, we generated a sample path for eqn. (7.16) and plot three variables in \mathbf{X} : the voltage V , Na^+ channel open probability M_{31} and K^+ channel open probability N_4 . The corresponding trajectory, \mathbf{X}^* , for Fox and Lu's model was generated from eqn. (7.17) and the corresponding Wiener trajectory was calculated using eqn. (7.18). To see that these trajectories are pathwise equivalent, in the top three subplots in panel A (resp. B) we superposed the voltage V^* , Na^+ channel open probability M_{31}^* and K^+ channel open probability N_4^* in \mathbf{X}^* against those in \mathbf{X} . The bottom three subplots in panel A (resp. B) plot the point-wise differences of each variable. Eqns. (7.16) and (7.17) are numerically solved in Matlab using the Euler-Maruyama method with a time step $dt = 0.001\text{ms}$. The slight differences observed arise in part due to numerical errors in calculating the singular value decomposition of \mathcal{S} (in eqn. (7.16)); another source of error is the finite accuracy of the Euler-Maruyama method.¹ As shown in Fig. 7.1, most differences occur near the spiking region, where the system is numerically very stiff and the numerical accuracy of

¹The forward Euler method is first order accurate for ordinary differential equations, but the forward Euler-Maruyama method is only $O(\sqrt{dt})$ accurate for stochastic differential equations [68].

the SDE solver accounts for most of the discrepancies (analysis of which is beyond the scope of this thesis). We can conclude from the comparison in Fig. 7.1 that the $14 \times 28\text{D}$ Langevin model is pathwise equivalent with the Fox and Lu’s model. Similarly, the same analogy applies for other channel-based Langevin models such that with the same diffusion matrix $\mathcal{D}(\mathbf{X})$.

We have shown that our “ $14 \times 28\text{D}$ ” model, with a 14-dimensional state space and 28 independent noise sources (one for each directed edge) is pathwise equivalent to Fox and Lu’s original 1994 model as well as other channel-based models (under corresponding boundary conditions) including [41, 48, 49, 90, 93]. As we shall see in §8, the pathwise equivalent models give statistically indistinguishable interspike interval distributions under the same BCs. We emphasize the importance of boundary conditions for pathwise equivalence. Two simulation algorithms with the same A_i and S_i matrices will generally have *nonequivalent* trajectories if different boundary conditions are imposed. For example, [21] employs the same dynamics as [90] away from the boundary, where ion channel state occupancy approaches zero or one. But where the latter allow trajectories to move freely across this boundary (which leads only to small, short-lived excursions into “nonphysical” values), Dangerfield imposes reflecting boundary conditions through a projection step at the boundary. As we will see below (§8), this difference in boundary conditions leads to a statistically significant difference in the ISI distribution, as well as a loss of accuracy when compared with the “gold standard” Markov chain simulation.

A: From 14D to Fox and Lu



B: From Fox and Lu to 14D

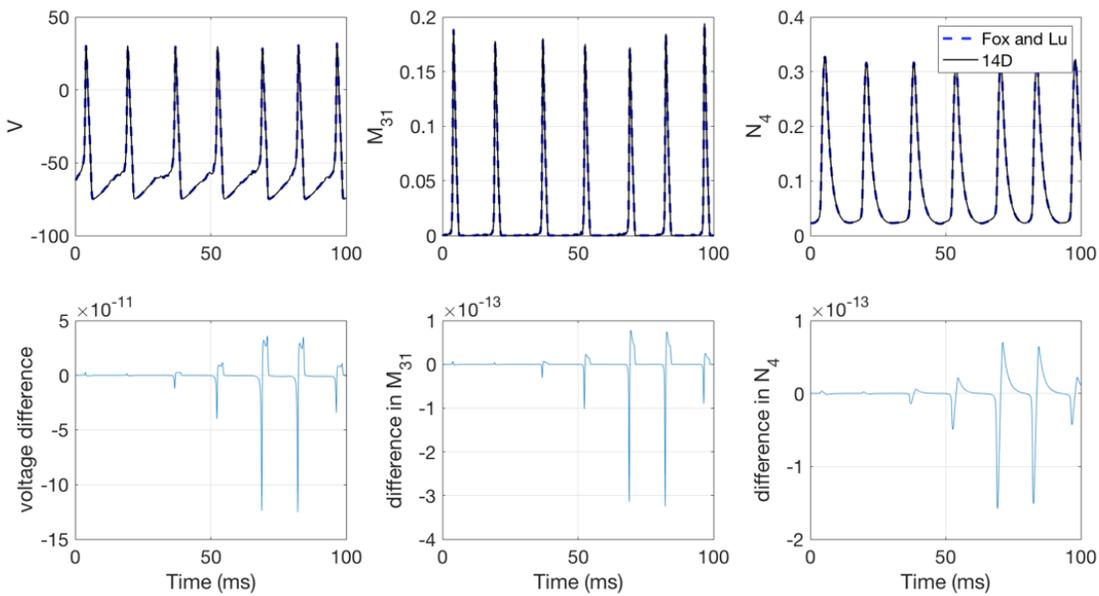


Figure 7.1: Pathwise equivalency of 14D HH model and Fox and Lu’s model. **A:** Given a sample path of the 14×28 D Langevin model in eqn. (7.16), we construct the noise by eqn. (7.18) and generate the sample trajectory of Fox and Lu’s model using eqn. (7.17). **B:** Given a sample path of Fox and Lu’s model in eqn. (7.17), we construct the noise by eqn. (7.24) and generate the sample trajectory of the 14×28 D Langevin model using eqn. (7.16). We obtain excellent agreement in both directions. See text for details.

Chapter 8

Model Comparison

In §6, we studied the contribution of every directed edge to the ISI variability and proposed how stochastic shielding could be applied under current clamp. Moreover, in §7, we proved that a family of Langevin models are pathwise equivalent.

Here we compare the accuracy and computational efficiency of several models, including the “subunit model” [40, 49], Langevin models with different S matrices or boundary conditions [21, 39, 49, 90, 93], the 14D HH model (proposed in §6.3), the 14D stochastic shielding model with six independent noise sources (proposed in §6.5), and the “gold standard” Markov Chain model (discussed in §6.2). Where other studies have compared moment statistics such as the mean firing frequency (under current clamp) and stationary channel open probabilities (under voltage clamp), we base our comparison on the entire interspike interval (ISI) distributions, under current clamp with a common fixed driving current. We use two different comparisons of ISI distributions, the first based on the L_1 norm of the difference between two distributions (the Wasserstein distance, [124]), in §8.1 and the second based on the L_∞ norm (the Kolmogorov-Smirnov test, [69, 117]), in §8.2.

We find similar results using both measures: as expected, the models that produce pathwise equivalent trajectories (Fox & Lu '94, Orio & Soudry, and our $14 \times 28\text{D}$ model) have indistinguishable ISI statistics, while the non-equivalent models (Fox '97, Dangerfield, Goldwyn & Shea-Brown, our $14 \times 6\text{D}$ stochastic-shielding model) have significantly different ISI distributions. Of these, the $14 \times 6\text{D}$ SS model is the closest to the models in the $14 \times 28\text{D}$ class, and as fast as any other model considered.

8.1 L_1 Norm Difference of ISIs

We first evaluate the accuracy of different stochastic simulation algorithms by comparing their ISI distributions under current clamp to that produced by a reference algorithm, namely the discrete-state Markov Chain (MC) algorithm.

Let X_1, X_2, \dots, X_n be n independent samples of ISIs with a true cumulative distribution function F . Let $F_n(\cdot)$ denote the corresponding empirical cumulative distribution function (ECDF) defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}, \quad x \in \mathbb{R}, \quad (8.1)$$

where we write $\mathbf{1}_A$ to denote the indicator function for the set A . Let Q and Q^M be the quantile functions of F and F^M , respectively. The L_1 -Wasserstein distance between two CDF's F^M and F can be written as [114] (page 64)

$$\rho_1(F, F^M) = \int_0^\infty |F(x) - F^M(x)| dx = \int_0^1 |Q(x) - Q^M(x)| dx. \quad (8.2)$$

Note that ρ_1 has the same units as “ dx ”. Thus the L_1 distances reported in Tab. 8.1 have units of milliseconds.

When two models have the same number of samples, n , (8.2) can be estimated by

$$\int_0^1 |Q(x) - Q^M(x)| dx \approx \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| := \rho_1(F_n, F_n^M), \quad (8.3)$$

where X_1, \dots, X_n and Y_1, \dots, Y_n are n independent samples sorted in ascending order with CDF F and F^M , respectively.

We numerically calculate $\rho_1(F_n, F_n^M)$ to compare several Langevin models against the MC model. We consider the following models: “Fox94” denotes the original model proposed by [39], which requires a square root decomposition ($S = \sqrt{D}$) for each step in the simulation, see equations (3.1)-(3.3). “Fox97” is the widely used “subunit model” of [40], see equations (3.4)-(3.5). “Goldwyn” denotes the method taken from [49], where they restrict the 14D system (V , 5 K^+ gates and 8 Na^+ gates) to the 4D multinomial submanifold (V , m , n , and h , see p. 35 above), with gating variables truncated to $[0, 1]$. We write “Orio” for the model proposed by [90], where they constructed a rectangular matrix S such that $SS^T = D$ (referred to as S_{paired} in Tab. 8.1) combining fluctuations driven by pairs of reciprocal edges, thereby avoiding taking matrix square roots at each time step. The model “Dangerfield” represents [21], which used the same S matrix as in [90] but added a reflecting (no-flux) boundary condition via orthogonal projection (referred to as S_{EF} in Tab. 8.1). Finally, we include the 14×28 D model we proposed in §6.3, or “14D” (referred to as S_{single} in Tab. 8.1); “SS” is the stochastic shielding model specified in §6.5.

For each model, we ran 10,000 independent samples of the simulation, holding channel number, injected current ($I_{\text{app}} = 10$ nA), and initial conditions fixed. Throughout the thesis, we presume a fixed channel density of 60 channels/ μm^2 for sodium and 18 channels/ μm^2 for potassium in a membrane patch of area $100 \mu\text{m}^2$, consistent with prior work such as [49, 90]. The initial condition is taken to be the point on the deterministic limit cycle at

which the voltage crosses upwards through -60 mV. An initial transient corresponding to 10-15 ISIs is discarded, to remove the effects of the initial condition. See Tab. B.1 in Appendix B for a complete specification of simulation parameters. We compared the efficiency and accuracy of each model through the following steps:

1. For each model, a single run simulates a total time of 84000 milliseconds (ms) with time step 0.008 ms, recording at least 5000 ISIs.
2. For each model, repeat 10,000 runs in step one.
3. Create a reference ISI distribution by aggregating all 10,000 runs of the MC model, i.e. based on roughly 5×10^7 ISIs.
4. For each of 10^4 individual runs, align all ISI data into a single vector and calculate the ECDF using equation (8.1).
5. Compare the ISI distribution of each model with the reference MC distribution by calculating the L_1 -difference of the ECDFs using equation (8.3).
6. To compare the computational efficiency, we take the average execution time of the MC model as the reference. The relative computational efficiency is the ratio of the average execution time of a model with that of the MC model (c. 3790 sec.).

Table 8.1 gives the empirically measured L_1 difference in ISI distribution between several pairs of models.¹ The first row (“MC”) gives the average L_1 distance between individual MC simulations and the reference distribution generated by aggregating all MC simulations, in order to give an estimate of the intrinsic variability of the measure.

¹Runtimes in Tab. 8.1, rounded to the nearest integer number of seconds, were obtained by averaging the runtimes on a distribution of heterogeneous compute nodes from Case Western Reserve University’s high-performance computing cluster.

Model	Variables V+M+N	S Matrix	Noise Dim. Na+K	L_1 Norm (msec.) (Wasserstein Dist.)	Runtime (sec.)
MC	1+8+5	n/a	20+8	$2.27 \text{ e-}4 \pm 7.15 \text{ e-}5$	3790
Fox94	1+7+4	$S = \sqrt{D}$	7+4	$4.74 \text{ e-}2 \pm 1.93 \text{ e-}4$	2436
Fox97	1+2+1	n/a	3	$8.01 \text{ e-}1 \pm 9.48 \text{ e-}4$	67
Dangerfield	1+8+5	S_{EF}	10+4	$2.18 \text{ e-}1 \pm 2.14 \text{ e-}4$	655
Goldwyn	1+8+5	$S = \sqrt{D}$	8+5	$1.83 \text{ e-}1 \pm 1.93 \text{ e-}4$	2363
Orio	1+8+5	S_{paired}	10+4	$4.52 \text{ e-}2 \pm 2.08 \text{ e-}4$	577
$14 \times 28\text{D}$	1+8+5	S_{single}	20+8	$4.93 \text{ e-}2 \pm 1.94 \text{ e-}4$	605
SS	1+8+5	S_{ss}	4+2	$7.62 \text{ e-}2 \pm 7.57 \text{ e-}5$	73

Table 8.1: Summary of the L_1 -Wasserstein distances of ISI distributions for Langevin type Hodgkin-Huxley models compared to the MC model. Model (see text for details): MC: Markov-chain. Fox1994: model from [39]. Fox97: [40]. Goldwyn: [49]. Dangerfield: [21, 20]. $14 \times 28\text{D}$: model proposed in §6.3. SS: stochastic-shielding model (§6.5). Variables: number of degrees of freedom in Langevin equation representing voltage, sodium gates, and potassium gates, respectively. S Matrix: Form of the noise coefficient matrix in equations (3.1)-(3.3). Noise Dimensions: number of independent Gaussian white-noise sources represented for sodium and potassium, respectively. L_1 Norm: Empirically estimated L_1 -Wasserstein distance between the model’s ISI distribution and the MC model’s ISI distribution. For MC-vs-MC, independent trials were compared. $a \pm b$: mean \pm standard deviation. Runtime (in sec.): see text for details. n/a: not applicable.

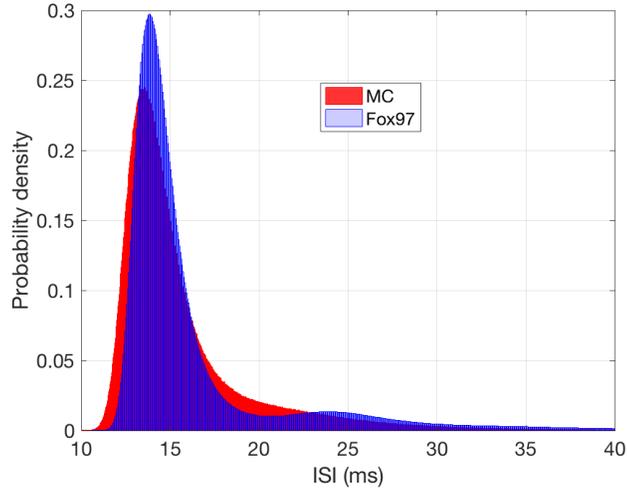


Figure 8.1: The probability density of interspike intervals (ISIs) for Fox97 (blue) and the MC model (red). The probability densities were calculated over more than 5.4×10^7 ISIs.

Figure 8.2 plots the L_1 -Wasserstein differences versus the relative computational efficiency of several models against the MC model. These results suggest that the Fox94, Orio, and $14 \times 28D$ models are statistically indistinguishable, when compared with the MC model using the L_1 -Wasserstein distance. This result is expected in light of our results (§7) showing that these three models are pathwise-equivalent. (We will make pairwise statistical comparisons between the ISI distributions of each model in §8.2.) Among these equivalent models, however, the $14 \times 28D$ and Orio models are significantly faster than the original Fox94 model (and the Goldwyn model) because they avoid the matrix square root computation. The Dangerfield model has speed similar to the $14 \times 28D$ model, but the use of reflecting boundary conditions introduces significant inaccuracy in the ISI distribution. The imposition of truncating boundary conditions in the Goldwyn model also appears to affect the ISI distribution. Of the models considered, the Fox97 subunit model is the fastest, however it makes a particularly poor approximation to the ISI distribution of the MC model. Note that the maximum L_1 -Wasserstein distance between two distributions is 2. The ISI

distribution of Fox97 subunit model to that of the MC model is more than 0.8, which is ten times larger than the L_1 -Wasserstein distance of the SS model, and almost half of the maximum distance. As shown in Fig. 8.1, the Fox97 subunit model fails to achieve the spike firing threshold and produces longer ISIs. Because of its inaccuracy, we do not include the subunit model in our remaining comparisons. The stochastic shielding model, on the other hand, has nearly the same speed as the Fox97 model, but is over 100 times more accurate (in the L_1 sense). The SS model is an order of magnitude faster than the 14×28 D model, and has less than twice the L_1 discrepancy versus the MC model (L_1 norm 76.2 versus 49.3 microseconds). While this difference in accuracy is statistically significant, it may not be practically significant, depending on the application (see §12 for further discussion of this point).

8.2 Two-sample Kolmogorov-Smirnov Test

In addition to using the L_1 -Wasserstein distances to test the differences between two CDFs, we can also make a pairwise comparison between each model by applying the Dvoretzky-Kiefer-Wolfowitz inequality [28] and the two-sample Kolmogorov-Smirnov (KS) test [69, 117]. While the Wasserstein distance is based on the L_1 norm, the KS statistic is based on the L_∞ (or supremum) norm.

The Dvoretzky-Kiefer-Wolfowitz inequality [28] establishes confidence bounds for the CDF. Specifically, the interval that contains the true CDF, $F(\cdot)$, with probability $1 - \alpha$, is given by

$$|F_n(x) - F(x)| \leq \varepsilon \quad \text{where } \varepsilon = \sqrt{\frac{\ln \frac{2}{\alpha}}{2n}}. \quad (8.4)$$

When comparing samples $X_1^M, X_2^M, \dots, X_n^M$ obtained from an approximate model M

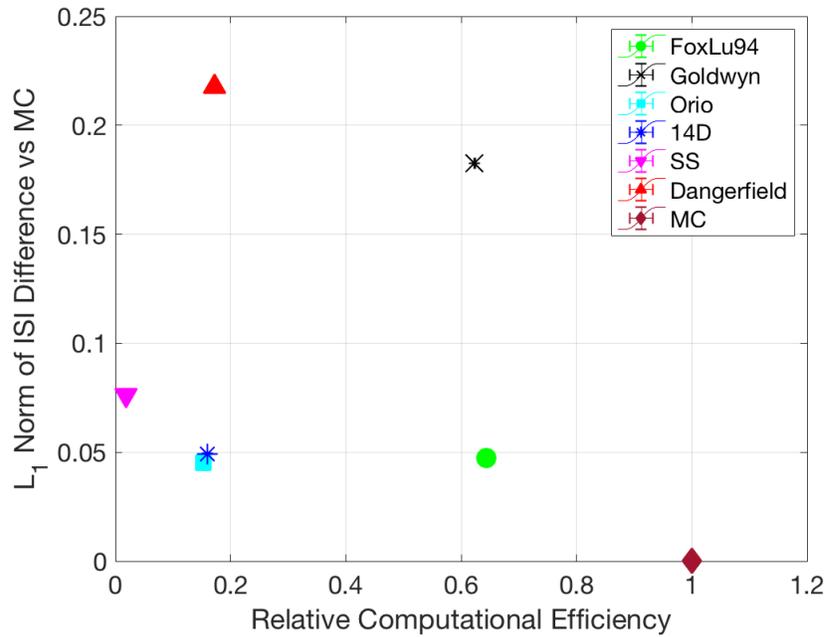


Figure 8.2: The L_1 -Wasserstein distances and relative computational efficiency vs. the MC model. “Fox94” (green circle), “Goldwyn” (black cross), “Orio” (cyan square), “14D” (blue star), “SS” (magenta downward pointing triangle), “Dangerfield” (red upward pointing triangle), and the “MC” (brown diamond) model. The L_1 error for ISI distribution was computed using the L_1 -Wasserstein distance (8.3), with discrete time Gillespie/Monte-Carlo simulations as a reference. The relative computational efficiency is the ratio of the recorded run time to the mean recorded time of the MC mode (3790 seconds). The mean and 95% confidence intervals were calculated using 100 repetitions of 10,000 runs each (5×10^9 ISIs total).

against the gold standard, in §8.1 we computed the L_1 difference of the empirical density functions, as an approximation for the L_1 difference of the true distributions. Instead, we work here with the L_∞ norm,

$$\rho_\infty(F_n, F_n^M) = \lim_{p \rightarrow \infty} \left(\int_0^\infty |F_n^M(x) - F_n(x)|^p dx \right)^{1/p} = \sup_{0 \leq x < \infty} (|F_n^M(x) - F_n(x)|). \quad (8.5)$$

For each $x \geq 0$, equation (8.4) bounds the discrepancy between the true and empirical distribution differences as follows. By the triangle inequality, and independence of the X_i from the X_i^M , the inequality

$$\begin{aligned} |F^M - F| &= |F^M - F_n^M + F_n - F + F_n^M - F_n| \\ &\leq |F^M - F_n^M| + |F_n - F| + |F_n^M - F_n| \\ &\leq 2\varepsilon + |F_n^M - F_n|, \end{aligned} \quad (8.6)$$

holds with probability $(1 - \alpha)^2$. Similarly,

$$\begin{aligned} |F_n^M - F_n| &= |F_n^M - F_n^M + F - F_n + F^M - F| \\ &\leq |F^M - F_n^M| + |F_n - F| + |F^M - F| \\ &\leq 2\varepsilon + |F^M - F| \end{aligned} \quad (8.7)$$

also holds with probability $(1 - \alpha)^2$. Together, (8.6)-(8.7) indicate that the discrepancy between the difference of empirical distributions and the difference of true distributions is bounded as

$$\left| |F^M - F| - |F_n^M - F_n| \right| \leq 2\varepsilon \quad (8.8)$$

with probability $(1 - \alpha)^2$, for $\varepsilon = \sqrt{\frac{\ln \frac{2}{\alpha}}{2n}}$.

We will use the pointwise difference of the ECDF's for a large sample as an estimate for the pointwise difference between two true CDFs. The two-sample Kolmogorov-Smirnov (KS) test [69, 117] offers a statistics to test whether two samples are from the same distribution. The two-sample KS statistic is

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|, \quad (8.9)$$

where $F_{1,n}$ and $F_{2,m}$ are two ECDFs for two samples defined in (8.1), and the sup is the supremum function. The reference statistic, $R_{n,m}(\alpha)$, depending on the significance level α , is defined as

$$R_{n,m}(\alpha) = \sqrt{\frac{-\log(\alpha/2)}{2}} \sqrt{\frac{n+m}{nm}}, \quad (8.10)$$

where n and m are the sample sizes. The null hypothesis that “the two samples come from the same distribution” is rejected at the significance level α if

$$D_{n,m} > R_{n,m}(\alpha). \quad (8.11)$$

Figure 8.3 plots the logarithm of ratio of the two-sample KS statistics, $\frac{D_{n,m}}{R_{n,m}(0.01)}$, for “Fox94” [39], “Goldwyn” [49], “Dangerfiled” [21], “Orio” [90], “14D” (the $14 \times 28D$ model we proposed in §6.3). Data of “self-comparison” (e.g. Fox94 vs. Fox 94) was obtained by comparing two ISI ECDF's from independent simulations. As shown in Fig. 8.3, models that we previously proved were pathwise equivalent in §7, namely the “Fox94”, “Orio” and the “14D” model, are not distinguishable at any confidence level justified by our data. Note that those three models use the same boundary conditions (free boundary condition as in [90]) and the ratio $D_{n,m}/R_{n,m}(\alpha)$ of pairwise comparison is on the same magnitude of that for the self-comparisons. However, as pointed out above, these statistically equivalent

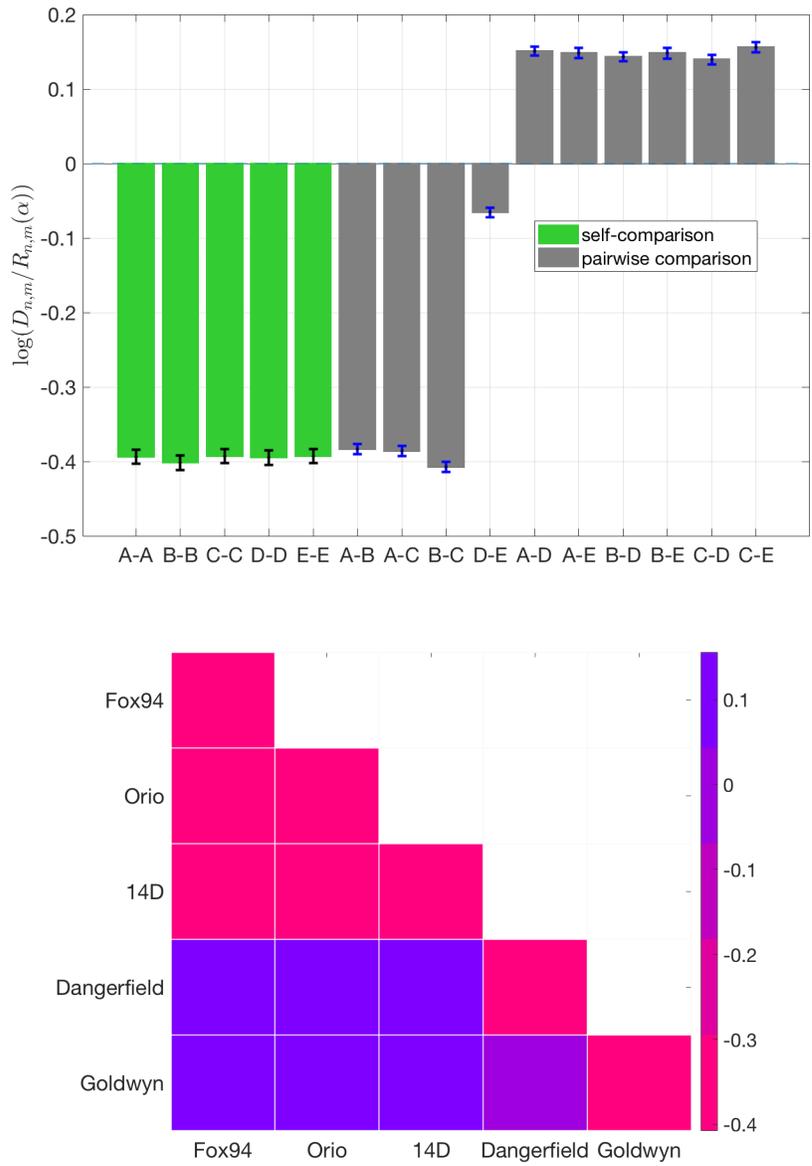


Figure 8.3: Logarithm of the ratio of Kolmogorov-Smirnov test statistic $D_{n,m}/R_{n,m}(\alpha)$, eqns. (8.9)-(8.10), for samples from the ISI distribution for each pair of models. **Top:** Box and whisker plots showing mean and 95% confidence intervals based on 10,000 pairwise comparisons. The first five plots show self-comparisons (green bars); the remainder compare distinct pairs (grey bars). A:“Fox94” [39], B:“Orio” [90], C: “14D” (14 × 28D model §6.3), D: Dangerfield [21], E: Goldwyn [49]. **Bottom:** Mean logarithms (as in top panel) for all pairwise comparisons.

simulation algorithms have different computational efficiencies (Fig. 8.2). Among these methods, Orio and Sundry’s algorithm (14 dimensional state space with 14 undirected edges as noise sources) and our method (14 dimensional state space with 28 directed edges as noise sources) have similar efficiencies, with Orio’s method being about 5% faster than ours method. Our $14 \times 28D$ method provides the additional advantage that it facilitates further acceleration under the stochastic shielding approximation (see §12).

In contrast to the statistically equivalent Orio, $14 \times 28D$ and Fox ’94 models, algorithms using different boundary conditions are not pathwise equivalent, which is again verified in Fig. 8.3. Algorithms with subunit approximation and truncated boundary condition (i.e., “Goldwyn”) and the reflecting boundary condition (i.e. “Dangerfield”) are significantly different in accuracy (and in particular, they are less accurate) than models in the $14 \times 28D$ class.

Part IV

Applications of the $14 \times 28D$ Langevin Model

Chapter 9

Definitions, Notations and Terminology

In this chapter, we will present definitions, notations and terminology that are necessary for the main applications of the $14 \times 28\text{D}$ HH model. We adopt the standard convention that uppercase symbols (e.g. $V, \mathbf{M}, \mathbf{N}$) represent random variables, while corresponding lowercase symbols (e.g. $v, \mathbf{m}, \mathbf{n}$) represent possible realizations of the random variables. Thus $\mathbb{P}(V \leq v)$ is the probability that the random voltage V does not exceed the value v . We set vectors in bold font and scalars in standard font.

9.1 The HH Domain

Recall the $14 \times 28\text{D}$ Langevin model

$$C \frac{dV}{dt} = I_{\text{app}}(t) - \bar{g}_{\text{Na}} \mathbf{M}_8 (V - V_{\text{Na}}) - \bar{g}_{\text{K}} \mathbf{N}_5 (V - V_{\text{K}}) - g_{\text{leak}} (V - V_{\text{leak}}), \quad (9.1)$$

$$\frac{d\mathbf{M}}{dt} = A_{\text{Na}} \mathbf{M} + S_{\text{Na}} \xi_{\text{Na}}, \quad (9.2)$$

$$\frac{d\mathbf{N}}{dt} = A_{\text{K}} \mathbf{M} + S_{\text{K}} \xi_{\text{K}}. \quad (9.3)$$

The drift matrices A_{Na} and A_{K} , S_{Na} and S_{K} are given in §6.3, $\xi_{\text{Na}} \in \mathbb{R}^{20}$ and $\xi_{\text{K}} \in \mathbb{R}^8$ are vectors of independent Gaussian white noise processes with zero mean and unit variance. Reversal potentials V_{ion} for physiological ions are typically confined to the range $\pm 150\text{mV}$. For the 4-D and the 14-D HH models, the reversal potentials for K^+ and Na^+ are -77mV and $+50\text{mV}$ respectively [32]. In Lemma 3, we establish that the voltage for conductance-based model in eqn. (9.1)-(9.3) is bounded. Therefore we can find a voltage range $[v_{\min}, v_{\max}]$ that is forward invariant with probability 1, meaning that the probability of any sample path leaving the range $v_{\min} \leq V(t) \leq v_{\max}$ is zero. At the same time, the channel state distribution for any channel with k states is restricted to a $(k-1)$ -dimensional simplex $\Delta^{k-1} \subset \mathbb{R}^k$, given by $y_1 + \dots + y_k = 1, y_i \geq 0$. Therefore, the phase space of any conductance-based model of the form eqn. (9.1)-(9.3) may be restricted to a compact domain in finite dimensions.

Definition 1. We define the HH domain \mathcal{D} to be

$$\mathcal{D} \triangleq [v_{\min}, v_{\max}] \times \Delta^7 \times \Delta^4 \quad (9.4)$$

where Δ^7 is the simplex occupied by the Na^+ channel states, and Δ^4 is occupied by the K^+ channel states.

We thus represent the “14-D” HH model in a reduced phase space of dimension $1+7+4=12$.

Lemma 3. For a conductance-based model of the form in eqn. (9.1)-(9.3), and for any bounded applied current $I_- \leq I_{\text{app}} \leq I_+$, there exist upper and lower bounds v_{\max} and v_{\min} such that trajectories with initial voltage condition $v \in [v_{\min}, v_{\max}]$ remain within this interval for all times $t > 0$, with probability 1, regardless of the initial channel state.

Proof. Let $V_1 = \min_{\text{ion}} \{V_{\text{ion}}\} \wedge V_{\text{leak}}$, and $V_2 = \max_{\text{ion}} \{V_{\text{ion}}\} \vee V_{\text{leak}}$, where $\text{ion} \in \{\text{Na}^+, \text{K}^+\}$. Note that for both the Na^+ and K^+ channel, $0 \leq \mathbf{M}_8 \leq 1$, $0 \leq \mathbf{N}_5 \leq 1$, with probability

1. Moreover, $g_i > 0$, $g_{\text{leak}} > 0$, therefore when $V \leq V_1$

$$\frac{dV}{dt} = \frac{1}{C} \{I_{\text{app}}(t) - \bar{g}_{\text{Na}}\mathbf{M}_8 (V - V_{\text{Na}}) - \bar{g}_{\text{K}}\mathbf{N}_5 (V - V_{\text{K}}) - g_{\text{leak}}(V - V_{\text{leak}})\} \quad (9.5)$$

$$\geq \frac{1}{C} \{I_{\text{app}}(t) - \bar{g}_{\text{Na}}\mathbf{M}_8 (V - V_1) - \bar{g}_{\text{K}}\mathbf{N}_5 (V - V_1) - g_{\text{leak}}(V - V_1)\} \quad (9.6)$$

$$\geq \frac{1}{C} \{I_{\text{app}}(t) - 0 \times \mathbf{M}_8 (V - V_1) - 0 \times \mathbf{N}_5 (V - V_1) - g_{\text{leak}}(V - V_1)\} \quad (9.7)$$

$$= \frac{1}{C} \{I_{\text{app}}(t) - g_{\text{leak}}(V - V_1)\}. \quad (9.8)$$

Inequality (9.6) holds with probability 1 because $V_1 = \min_{i \in \mathcal{I}} \{V_i\} \wedge V_{\text{leak}}$, and inequality (9.7) follows because $V - V_1 \leq 0$, $g_i > 0$ and $\mathbf{M}_8 \geq 0$, $\mathbf{N}_5 \geq 0$ with probability 1. Let $V_{\min} := \min \left\{ \frac{I_{\text{app}}}{g_{\text{leak}}} + V_1, V_1 \right\}$. When $V < V_{\min}$, $\frac{dV}{dt} > 0$ with probability 1. Therefore, V will not decrease beyond V_{\min} with probability 1.

Similarly, when $V \geq V_2$

$$\frac{dV}{dt} = \frac{1}{C} \{I_{\text{app}}(t) - \bar{g}_{\text{Na}}\mathbf{M}_8 (V - V_{\text{Na}}) - \bar{g}_{\text{K}}\mathbf{N}_5 (V - V_{\text{K}}) - g_{\text{leak}}(V - V_{\text{leak}})\} \quad (9.9)$$

$$\leq \frac{1}{C} \{I_{\text{app}}(t) - \bar{g}_{\text{Na}}\mathbf{M}_8 (V - V_2) - \bar{g}_{\text{K}}\mathbf{N}_5 (V - V_2) - g_{\text{leak}}(V - V_2)\} \quad (9.10)$$

$$\leq \frac{1}{C} \{I_{\text{app}}(t) - 0 \times \mathbf{M}_8 (V - V_2) - 0 \times \mathbf{N}_5 (V - V_2) - g_{\text{leak}}(V - V_2)\} \quad (9.11)$$

$$= \frac{1}{C} \{I_{\text{app}}(t) - g_{\text{leak}}(V - V_2)\}. \quad (9.12)$$

Inequality (9.10) and inequality (9.11) holds with probability 1 because $V_2 = \max_{i \in \mathcal{I}} \{V_i\} \vee V_{\text{leak}}$, $V - V_2 \geq 0$, $g_i > 0$ and $\mathbf{M}_8 \geq 0$, $\mathbf{N}_5 \geq 0$ with probability 1. Let $V_{\max} = \max \left\{ \frac{I_{\text{app}}}{g_{\text{leak}}} + V_2, V_2 \right\}$. When $V > V_{\max}$, $\frac{dV}{dt} < 0$ with probability 1. Therefore, V will not go beyond V_{\max} with probability 1.

We conclude that if V takes an initial condition in the interval $[V_{\min}, V_{\max}]$, then $V(t)$ remains within this interval for all $t \geq 0$ with probability 1. \square

Throughout this thesis, we use the parameters of the standard Hodgkin-Huxley model, with constant applied current $I_{\text{app}} = 10$ nano Amperes (see Tab. B.1). For these parameters, $V_{\text{min}} = -77$ mV and $V_{\text{max}} \approx +83.33$ mV.

Remark 3. *Lemma 3 implies that the per capita transition rates along a finite collection of edges, $\{\alpha_k(v)\}_{k=1}^K$ are bounded above by a constant α_{max} , as v ranges over $v_{\text{min}} \leq v \leq v_{\text{max}}$. This fact will help establish Theorem 4.*

9.2 Interspike Intervals and First Passage Times

Figure 9.1 shows a voltage trajectory generated by the 14-D stochastic HH model, under current clamp, with injected current in the range supporting steady firing. The regular periodicity of the deterministic model vanishes in this case. Nevertheless, the voltage generates spikes, which allows us to introduce a well defined series of spike times and interspike intervals (ISIs). For example, we may select a reference voltage such as $v_{\text{th}} = -20$ mV, with the property that within a neighborhood of this voltage, trajectories have strictly positive or strictly negative derivatives (dV/dt) with high probability.

In [100], they [suggested](#) that the stochastic (Langevin) 4-D HH model has a unique invariant stationary joint density for the voltage and gating variables, as well as producing a stationary point process of spike times. The ensemble of trajectories may be visualized by aligning the voltage spikes (Figure 9.1b), and illustrates that each trace is either rapidly increasing or else rapidly decreasing as it passes $v_{\text{th}} = -20$ mV.

In order to give a precise definition of the interspike interval, on which we can base a first-passage time analysis, we will consider two types of Poincaré section of the fourteen-

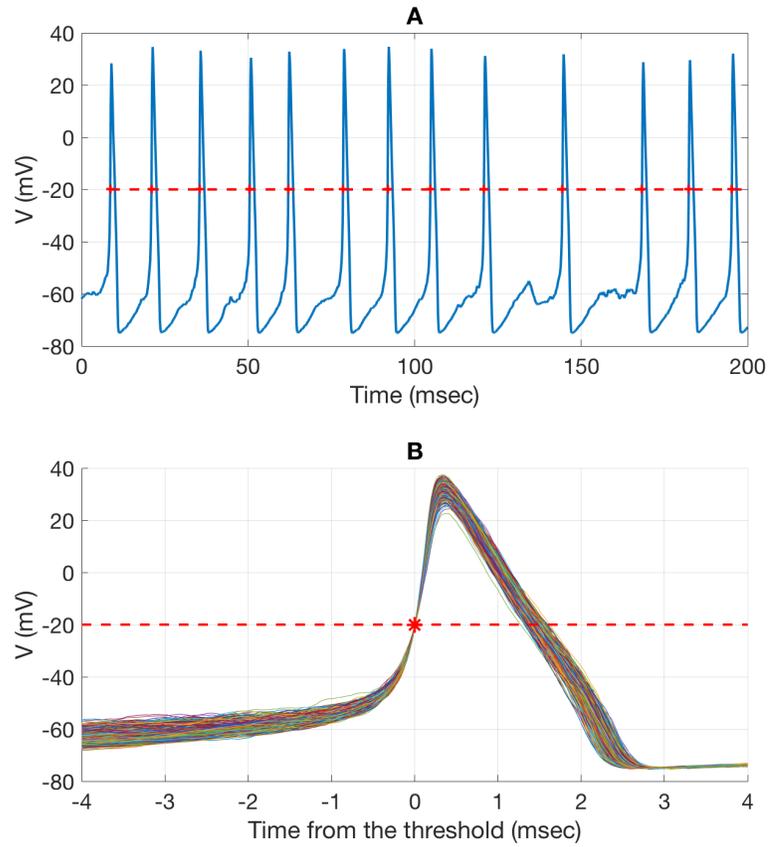


Figure 9.1: Example voltage trace for the 14-D stochastic HH model (eqn. (9.1)-(9.3)). **A** Voltage trace generated by the full 14-D stochastic HH model with $I_{app}=10$ nA; for other parameters see §B. The red dashed line indicates a threshold voltage at -20 mV, which is set at the rising phase of the spike. **B** Ensemble of voltage traces constructed by aligning traces with a voltage upcrossing at $v_{th} = -20$ mV (red star) for more than 1000 cycles.

dimensional phase space: the “nullcline” surface associated with the voltage variable,

$$\mathcal{V}^0 = \{(v, \mathbf{m}, \mathbf{n}) \in \mathcal{D} \mid f(v, \mathbf{m}, \mathbf{n}) = 0\}, \quad (9.13)$$

where the rate of change of voltage is instantaneously zero, and an iso-voltage sections of the form

$$\mathcal{S}^u = \{(v, \mathbf{m}, \mathbf{n}) \in \mathcal{D} \mid v = u\}. \quad (9.14)$$

(In §9.5 we will define a third type of Poincaré section, namely isochrons of the mean–return-time function $T(v, \mathbf{n})$ [75].) Figure 9.2 illustrates the projection of \mathcal{V}^0 (green horizontal line) and \mathcal{S}^u for $u \in \{-40, 10\}$ (red vertical lines) onto the $(V, dV/dt)$ plane.

For any voltage u we can partition the voltage-slice \mathcal{S}^u into three disjoint components $\mathcal{S}^u = \mathcal{S}_0^u \sqcup \mathcal{S}_+^u \sqcup \mathcal{S}_-^u$, defined as follows:

Definition 2. *Given the stochastic differential equations (6.42) defined on the HH domain \mathcal{D} , and for a given voltage u , the “null” surface, \mathcal{S}_0^u is defined as*

$$\mathcal{S}_0^u \triangleq \mathcal{S}^u \cap \mathcal{V}^0 = \{(v, \mathbf{m}, \mathbf{n}) \in \mathcal{D} \mid v = u \ \& \ f(v, \mathbf{m}, \mathbf{n}) = 0\},$$

the “inward current” surface, \mathcal{S}_+^u is defined as

$$\mathcal{S}_+^u \triangleq \{(v, \mathbf{m}, \mathbf{n}) \in \mathcal{D} \mid v = u \ \& \ f(v, \mathbf{m}, \mathbf{n}) > 0\},$$

and the “outward current” surface is defined as

$$\mathcal{S}_-^u \triangleq \{(v, \mathbf{m}, \mathbf{n}) \in \mathcal{D} \mid v = u \ \& \ f(v, \mathbf{m}, \mathbf{n}) < 0\}.$$

Figure 9.2 plots dV/dt versus V for roughly 1000 cycles, and shows that for certain

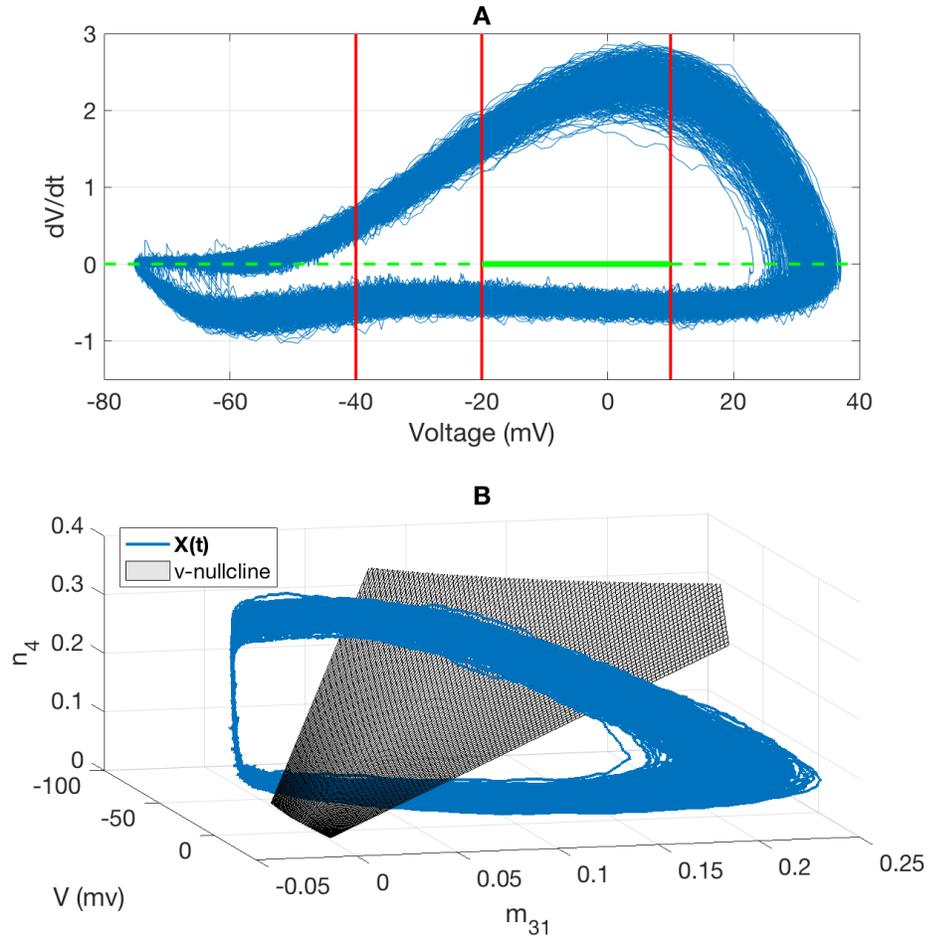


Figure 9.2: Sample trace of $\mathbf{X}(t)$ over 1000 cycles. **A:** View of $\mathbf{X}(t)$ (blue trace) in $(V, dV/dt)$ phase plane. Red vertical line: projections of voltage slices \mathcal{S}^{-40} , \mathcal{S}^{-20} and \mathcal{S}^{+10} . Green horizontal line: projection of V -nullcline \mathcal{V}^0 ; solid portion corresponds to voltage range in second panel. **B:** Projection of $\mathbf{X}(t)$ on three coordinates (V, m_{31}, n_4) . Gray surface: subset of v -nullcline with $-40 \text{ mV} \leq v \leq +10 \text{ mV}$. For this voltage range, trajectories remain a finite distance away from \mathcal{V}^0 with high probability.

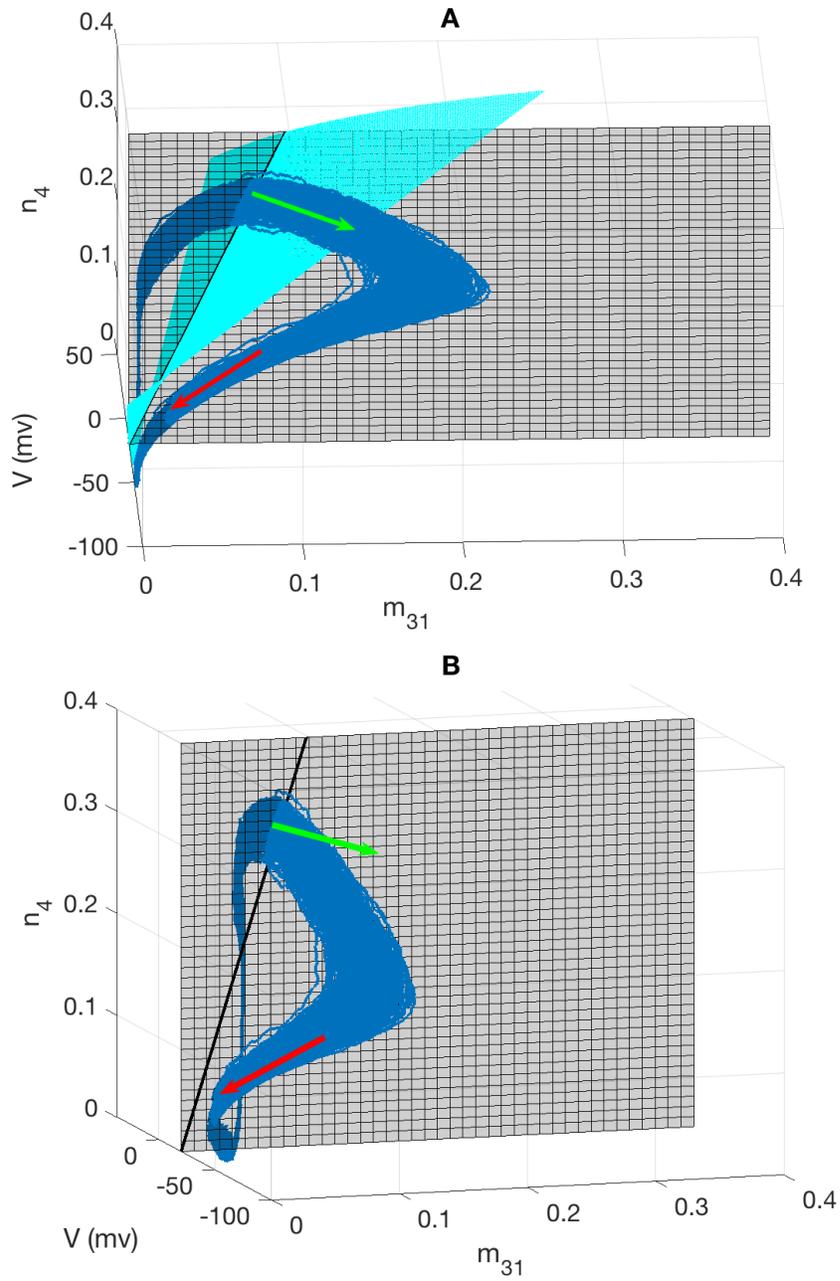


Figure 9.3: Intersections of a trajectory (blue trace) with a voltage slice (\mathcal{S}^{-20} , grey surface) and V -nullcline (\mathcal{V}^0 , cyan surface). **A**: Trajectory $\mathbf{X}(t)$ crosses \mathcal{S}_+^{-20} with increasing voltage component (green arrow). Subsequently, the trajectory crosses \mathcal{S}_-^{-20} with decreasing voltage component (red arrow). With probability one, the trajectory $\mathbf{X}(t)$ does not intersect with the null space for voltage in the range of $[-40, 10]$ mV. **B**: A special case for A with the null component \mathcal{S}_0^{-20} (black diagonal line) indicated for $v = -20$ mV. The intersection of the stochastic trajectory and $v = -20$ mV is partitioned into an inward component \mathcal{S}_+^{-20} (green arrow shows trajectory crossing with $dV/dt > 0$) and an outward component \mathcal{S}_-^{-20} (red arrow shows trajectory crossing with $dV/dt < 0$). Note that with probability one, the null component \mathcal{S}_0^{-20} does not intersect with the trajectory.

values of v , the density of trajectories in a neighborhood of \mathcal{V}^0 is very small for a finite voltage range (here shown as -40 to $+10$ mV). Indeed for any u , the intersection of the null set \mathcal{S}_0^u has measure zero relative to the uniform measure on \mathcal{S}^u , and the probability of finding a trajectory at precisely $V = u$ and $dV/dt = 0$ is zero. From this observation, and because dV/dt is conditionally deterministic, given \mathbf{n} , it follows that a trajectory starting from $\mathbf{x} \in \mathcal{S}_+^u$ will necessarily cross \mathcal{S}_-^u before crossing \mathcal{S}_+^u again (with probability one).

First-Passage Times Based on this observation, we can give a formal definition of the first passage time as follows.

Definition 3. *Given a section $\mathcal{S} \subset \mathcal{D}$, we define the first passage time (FPT) from a point $\mathbf{x} \in \mathcal{D}$ to \mathcal{S} , for a stochastic conductance-based model as*

$$\tau(\mathbf{x}, \mathcal{S}) \triangleq \inf\{t > 0 \mid \mathbf{X}(t) \in \mathcal{S} \ \& \ \mathbf{X}(0) = \mathbf{x}\}. \quad (9.15)$$

Note that, more generally, we can think of τ as $\tau(\mathbf{x}, \mathcal{S}, \omega)$, where ω is a sample from the underlying Langevin process sample space, $\omega \in \Omega$.¹ For economy of notation we usually suppress ω , and may also suppress \mathcal{S} , or \mathbf{x} when these are clear from context.

In the theory of stochastic processes a *stopping time*, τ , is any random time such that the event $\{\tau \leq t\}$ is part of the σ -algebra generated by the filtration \mathcal{F}_t of the stochastic process from time 0 through time t . That is, one can determine whether the event defining τ has occurred or not by observing the process for times up to and including t (see [89], §7.2, for further details).

Remark 4. *Given any section $\mathcal{S} \subset \mathcal{D}$ and any point $\mathbf{x} \in \mathcal{D}$, the first passage time $\tau(\mathbf{x}, \mathcal{S})$ is a stopping time. This fact will play a critical role in the proof of our main theorem.*

¹For the 14×28 D Langevin Hodgkin-Huxley model, Ω may be thought of as the space of continuous vector functions on $[0, \infty)$ with 28 components – one for each independent noise source.

As Figure 9.2 suggests, for $-40 \leq v \leq +10$ mV, the probability of finding trajectories in an open *neighborhood* of \mathcal{S}_0^v can be made arbitrarily small by making the neighborhood around \mathcal{S}_0^v sufficiently small. This observation has two important consequences. First, because the probability of being near the nullcline \mathcal{S}_0^v is vanishingly small, interspike intervals are well defined (cf. Def. 5, below), even for finite precision numerical simulation and trajectory analysis. In addition, this observation lets us surround the nullcline with a small cylindrical open set, through which trajectories are unlikely to pass. This cylinder-shaped buffer will play a role in defining the mean–return-time phase in §9.5.

Moreover, as illustrated in Figure 9.2, when $V = -20$ mV, the stochastic trajectory x intersects \mathcal{S}^{-20} at two points within each full cycle, where one is in \mathcal{S}_+^{-20} and one in \mathcal{S}_-^{-20} . In addition, the trajectory crosses \mathcal{S}_-^{-20} before it crosses \mathcal{S}_+^{-20} again. This is a particular feature for conductance-based models in which dV/dt is conditionally deterministic, i.e. the model includes no current noise.²

Definition 4. *Given any set $\mathcal{S} \subset \mathcal{D}$ (for instance, a voltage-section) and a point $\mathbf{x} \in \mathcal{D}$, the mean first passage time (MFPT) from \mathbf{x} to \mathcal{S} ,*

$$T(\mathbf{x}, \mathcal{S}) \triangleq \mathbb{E}[\tau(\mathbf{x}, \mathcal{S})], \quad (9.16)$$

and the second moment of the first passage time is defined as

$$S(\mathbf{x}, \mathcal{S}) \triangleq \mathbb{E}[\tau(\mathbf{x}, \mathcal{S})^2]. \quad (9.17)$$

²In this dissertation, we focus on a Langevin equation, rather than models with discrete channel noise. Therefore, our trajectories are diffusions, that have continuous sample paths (with probability one). Therefore, the FPT $\tau(\mathbf{x}, \mathcal{S})$ is well defined. For discrete channel-noise models, a slightly modified definition would be required.

Interspike Intervals Starting from $\mathbf{x} \in \mathcal{S}_+^{v_0}$, at time $t = 0$, we can identify the sequence of (τ, \mathbf{x}) pairs of crossing times and crossing locations as

$$(\tau_0^u = 0, \mathbf{x}_0^u = \mathbf{x}), (\tau_1^d, \mathbf{x}_1^d), (\tau_1^u, \mathbf{x}_1^u), \dots, (\tau_k^d, \mathbf{x}_k^d), (\tau_k^u, \mathbf{x}_k^u), \dots, \quad (9.18)$$

with $0 = \tau_0^u < \tau_1^d < \tau_1^u < \tau_2^d < \tau_2^u < \dots < \tau_k^d < \tau_k^u, \dots$

where $\tau_k^d = \inf\{t > \tau_{k-1}^u \mid \mathbf{x} \in \mathcal{S}_-^{v_0}\}$ is the k th down-crossing time, $\mathbf{x}_k^d \in \mathcal{S}_-^{v_0}$ is the k th down-crossing location, $\tau_k^u = \inf\{t > \tau_k^d \mid \mathbf{x} \in \mathcal{S}_+^{v_0}\}$ is the k th up-crossing time, and $\mathbf{x}_k^u \in \mathcal{S}_+^{v_0}$ is the k th up-crossing location, for all $k \in \mathbb{N}^+$.

Under constant applied current, the HH system has a unique stationary distribution with respect to which the sequence of crossing times and locations have well-defined probability distributions [100]. We define the moments of the interspike interval distribution with respect to this underlying stationary probability distribution.

Definition 5. *Given a sequence of up- and down-crossings, relative to a reference voltage v_0 as above, the k th interspike interval (ISI), I_k (in milliseconds), of the stochastic conductance-based model is a random variable that is defined as*

$$I_k \triangleq \tau_{k+1}^u - \tau_k^u \quad (9.19)$$

where τ_k^u is the k th up-crossing time. The mean ISI is defined as

$$I \triangleq \mathbb{E}[I_k] \quad (9.20)$$

and the second moment of the ISI is defined as

$$H \triangleq \mathbb{E}[I_k^2] \quad (9.21)$$

The variance of the ISI is defined as

$$\sigma_{ISI}^2 \triangleq \mathbb{E} [(I - I_k)^2], \quad (9.22)$$

where $k = 1, 2, \dots$.

It follows immediately that $\sigma_{ISI}^2 = H - I^2$.

9.3 Asymptotic phase and infinitesimal phase response curve

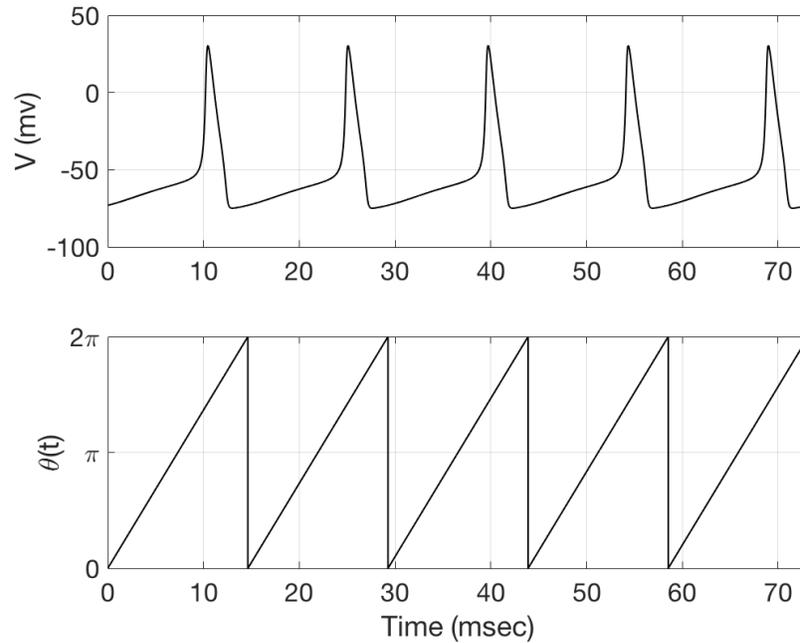


Figure 9.4: Sample trace on the limit cycle and its corresponding phase function $\theta(t)$. **Top:** voltage trace for the deterministic system $d\mathbf{x} = F(\mathbf{x}) dt$ showing a period of $T_0 \approx 14.63$ ms. **Bottom:** The phase function of time saceled from $[0, 2\pi)$.

Given parameters in App. B with an applied current $I_{app} = 10$ nA, the deterministic HH model,

$$\frac{d\mathbf{x}}{dt} = F(\mathbf{x}) \quad (9.23)$$

fires periodically with a period $T_0 \approx 14.63$ msec, as shown in Fig. 9.4. We assume that the deterministic model has an asymptotically stable limit cycle, $\gamma(t) = \gamma(t + T_0)$. The phase of the model at time t can be defined as [110]

$$\theta(t) = \frac{(t + T_0 \frac{\varphi}{2\pi}) \bmod T_0}{T_0} \times 2\pi, \quad (9.24)$$

where mod is the module operation, and $\theta(t) = 0$ sets the spike threshold for the model. The constant $\varphi \in [0, 2\pi]$ is the relative phase determined by the initial condition, and there is a one-to-one map between each point on the limit cycle and the phase. In general, the phase can be scaled to any constant interval; popular choices include $[0, 1)$, $[0, 2\pi)$, and $[0, T)$. Here we take $\theta \in [0, 2\pi)$ (see Fig. 9.4).

Winfree and Guckenheimer extended the definition of phase from the limit cycle to the basin of attraction, which laid the foundation for the asymptotic phase function $\phi(\mathbf{x})$ [54, 130, 131]. For the system in Eqn. (9.23), let $\mathbf{x}(0)$ and $\mathbf{y}(0)$ be two initial conditions, one on the limit cycle and one in the basin of attraction, respectively. Denote the phase associated to $\mathbf{x}(0)$ as $\theta_0(t)$. If the solutions $\mathbf{x}(t)$ and $\mathbf{y}(t)$ satisfy

$$\lim_{t \rightarrow \infty} |\mathbf{x}(t) - \mathbf{y}(t)| = 0, \text{ i.e. } \lim_{t \rightarrow \infty} |\phi(\mathbf{y}(t)) - \theta_0(t)| = 0,$$

then $\mathbf{y}(0)$ has asymptotic phase θ_0 . The set of all points sharing the same asymptotic phase comprises an *isochron*, a level set of $\phi(\mathbf{x})$. We also refer to such a set of points as an *isophase surface* [109]. By construction, the asymptotic phase function $\phi(\mathbf{x})$ coincides with the oscillator phase $\theta(t)$ on the limit cycle, i.e. $\theta(t) = \phi(\gamma(t))$. We will assume that $\phi(\mathbf{x})$ is twice-differentiable within the basin of attraction.

The phase response curve (PRC) is defined as the change in phase of an oscillating system in response to a given perturbations. If the original phase is defined as θ_b and the

phase after perturbation as θ_a , then the PRC is the shift in phase

$$\Delta(\theta) = \theta_a - \theta_b.$$

In the limit of small instantaneous perturbations, the PRC may be approximated by the *infinitesimal phase response curve* (iPRC) [110, 131]. For a deterministic limit cycle, the iPRC $\mathbf{Z}(t)$ obeys the adjoint equation [11]

$$\frac{d\mathbf{Z}}{dt} = -\mathcal{J}(\gamma(t))^\top \mathbf{Z}, \quad (9.25)$$

$$\mathbf{Z}(0) = \mathbf{Z}(T_0), \quad (9.26)$$

$$\mathbf{Z}(0)^\top \mathbf{F}(\gamma(0)) = 1 \quad (9.27)$$

where T_0 is the period of the deterministic limit cycle, $\gamma(t)$ is the periodic limit cycle trajectory (for the HH equations (9.23), $\gamma(t) \in \mathbb{R}^{14}$) and $\mathcal{J}(t) = D\mathbf{F}(\gamma(t))$ is the Jacobian of \mathbf{F} evaluated along the limit cycle. The iPRC $\mathbf{Z}(t)$ is proportional to the gradient of the phase function $\phi(\mathbf{x})$ evaluated on the limit cycle. For any point \mathbf{x} in the limit cycle's basin of attraction, we can define a timing sensitivity function $\tilde{\mathbf{Z}}(\mathbf{x}) \triangleq \frac{T_0}{2\pi} \nabla_{\mathbf{x}} \phi(\mathbf{x})$. For the limit cycle trajectory $\gamma(t)$, we have $\mathbf{Z}(t) = \tilde{\mathbf{Z}}(\gamma(t))$. The first component of \mathbf{Z} , for example, has units of msec/mv, or change in time per change in voltage.

9.4 Small-noise expansions

In the governing Langevin equation (6.42), the stochastic forcing components in $\mathcal{G} d\mathbf{W}$ are implicitly scaled by factors proportional to $1/\sqrt{N_{\text{ion}}}$, the number of sodium or potassium channels in the membrane. The larger the number of channels, the smaller the

noise. In their comparative study of different Langevin models, Goldwyn and Shea-Brown considered a patch of excitable membrane containing $M_{\text{ref}} = 6000$ sodium channels and $N_{\text{ref}} = 1800$ potassium channels [49].

Here, we make the dependence on system size (number of channels) explicit, by introducing a small parameter $\epsilon \propto N_{\text{ion}}^{-1}$. We therefore consider a one-parameter family of Langevin equations,

$$d\mathbf{X} = \mathbf{F}(\mathbf{X}) dt + \sqrt{\epsilon} \mathcal{G}(\mathbf{X}) d\mathbf{W}(t), \quad (9.28)$$

with effective numbers of $M_{\text{tot}} = M_{\text{ref}}/\epsilon$ sodium and $N_{\text{tot}} = N_{\text{ref}}/\epsilon$ potassium channels. For sufficiently small values of ϵ , the solutions remain close to the deterministic limit cycle; the (stochastic) interspike intervals will remain close to the deterministic limit cycle period \bar{T}_0 . If $\mathbf{X}(t)$ is a trajectory of (9.28), and $\phi(\mathbf{x})$ is any twice-differentiable function, then Ito's lemma [89] gives an expression for the increment of ϕ during a time increment dt , beginning from state \mathbf{X} :

$$d\phi(\mathbf{X}(t)) = (\nabla\phi(\mathbf{X})) \cdot d\mathbf{X} + \frac{\epsilon}{2} \sum_{ij} \frac{\partial^2 \phi(\mathbf{X})}{\partial x_i \partial x_j} dt \quad (9.29)$$

$$\begin{aligned} &= \left(\mathbf{F}(\mathbf{X}) \cdot \nabla\phi(\mathbf{X}) + \frac{\epsilon}{2} \sum_{ij} \frac{\partial^2 \phi(\mathbf{X})}{\partial x_i \partial x_j} \right) dt + \sqrt{\epsilon} (\nabla\phi(\mathbf{X}))^\top \mathcal{G}(\mathbf{X}) d\mathbf{W} \\ &= \mathcal{L}^\dagger[\phi(\mathbf{X})] dt + \sqrt{\epsilon} (\nabla\phi(\mathbf{X}))^\top \mathcal{G}(\mathbf{X}) d\mathbf{W} \end{aligned} \quad (9.30)$$

up to terms of order dt . The operator \mathcal{L}^\dagger defined by (9.29)-(9.30) is the formal adjoint of the Fokker-Planck or Kolmogorov operator [99], also known as the generator of the Markov process [89], or the Koopman operator [72].

Dynkin's formula, which we will use to prove our main result, is closely related to equation (9.30). Let $\mathbf{x} \in \mathcal{D}$ and let $\mathbb{E}^{\mathbf{x}}$ denote the probability law for the ensemble of

stochastic trajectories beginning at \mathbf{x} . Dynkin’s theorem ([89], §7.4) states that if ϕ is a twice-differentiable function on \mathcal{D} , and if τ is any stopping time (cf. Remark 4) such that $\mathbb{E}^{\mathbf{x}}[\tau] < \infty$, then

$$\mathbb{E}^{\mathbf{x}}[\phi(\mathbf{X}(\tau))] = \phi(\mathbf{x}) + \mathbb{E}^{\mathbf{x}} \left[\int_0^\tau \mathcal{L}^\dagger[\phi(\mathbf{X}(s))] ds \right]. \quad (9.31)$$

9.5 Iso-phase Sections

For the deterministic model, the isochrons form a system of Poincaré sections \mathcal{S}_φ , $\varphi \in [0, 2\pi]$, each with a constant return time equal to the oscillator period T_0 . When the system is perturbed by noise, $\epsilon > 0$ in (9.28), we consider a set of “iso-phase sections” based on a mean–return–time (MRT) construction, first proposed by [109] and rigorously analyzed by [75]. As shown in [75], the MRT iso-phase surfaces \mathcal{S} are the level sets of a function $T_\epsilon(\mathbf{x})$ satisfying the MRT property. Namely, if \mathcal{S} is an iso-phase section, then the mean time taken to return to \mathcal{S} , starting from any $\mathbf{x} \in \mathcal{S}$, after one full rotation, is equal to the mean period, \bar{T}_ϵ .

The construction in [75] requires that the Langevin equation (9.28) be defined on a domain with the topology of an n -dimensional cylinder, because finding the MRT function $T_\epsilon(\mathbf{x})$ involves specifying an arbitrary “cut” from the outer to the inner boundary of the cylinder. Conductance-based models in the steady-firing regime, where the mean-field equations support a stable limit cycle, can be well approximated by cylindrical domains. In particular, their variables are restricted to a compact range, and there is typically a “hole” through the domain in which trajectories are exceedingly unlikely to pass, at least for small noise.

As an example, consider the domain for the 14D HH equations (recall Defs. 1), namely

$\mathcal{D} \triangleq [v_{\min}, v_{\max}] \times \Delta^7 \times \Delta^4$. The p -dimensional simplex Δ^p is a bounded set, and, as established by Lemma 1, the trajectories of (9.28) remain within fixed voltage bounds with probability 1, so our HH system operates within a bounded subset of \mathbb{R}^{14} . To identify a “hole” through this domain, note that the set

$$\mathcal{S}_0^u \triangleq \mathcal{S}^u \cap \mathcal{V}^0 = \{(v, \mathbf{m}, \mathbf{n}) \in \mathcal{D} \mid v = u \ \& \ f(v, \mathbf{m}, \mathbf{n}) = 0\},$$

which is the intersection of the voltage nullcline \mathcal{V}^0 with the constant-voltage section \mathcal{S}^u , is rarely visited by trajectories under small noise conditions (Fig. 9.2B).

For $r > 0$, we define the open ball of radius r around \mathcal{S}_0^u as

$$\mathcal{B}_r(\mathcal{S}_0^u) \triangleq \left\{ \mathbf{x} \in \mathcal{D} \mid \min_{\mathbf{y} \in \mathcal{S}_0^u} (\|\mathbf{x} - \mathbf{y}\|) < r \right\}. \quad (9.32)$$

For the remainder of the thesis, we take the stochastic differential equation (9.28) to be defined on

$$\mathcal{D}_0 = \mathcal{D} \setminus \mathcal{B}_r(\mathcal{S}_0^u). \quad (9.33)$$

For sufficiently small $r > 0$, \mathcal{D}_0 is a space homeomorphic to a cylinder in \mathbb{R}^{14} . To see this, consider the annulus $\mathcal{A} = I_1 \times B^{13}$, where $I_1 = [0, 2\pi]$, and B^{13} is a simply connected subset of \mathbb{R}^{13} . That space is homotopy equivalent to a circle S^1 by contracting the closed interval parts to a point, and contracting the annulus part to its inner circle.

To complete the setup so that we can apply the theory from [75], we set boundary conditions $\sum_{ij} n_i (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_j T_\epsilon = 0$ at reflecting boundaries with outward normal \mathbf{n} on both the inner and outer boundaries of the cylinder. In addition, we choose an (arbitrary) section transverse to the cylinder, and impose a jump condition $T_\epsilon \rightarrow T_\epsilon + \bar{T}_\epsilon$ across this section, where \bar{T}_ϵ is mean oscillator period under noise level ϵ .

As showed in [75], this construction allows us to establish a well defined MRT function for a given noise level ϵ , $\bar{T}_\epsilon(\mathbf{x})$. We then obtain the iso-phase sections as level sets of $\bar{T}_\epsilon(\mathbf{x})$. We give a formal definition as follows.

Definition 6. *Given a fixed noise level $\epsilon \geq 0$, and an iso-phase surface \mathcal{S} for eqn. (9.28), we define the k th iso-phase interval (IPI) as the random variable*

$$\Delta_k \triangleq \mu_k - \mu_{k-1}, \quad (9.34)$$

where $\{\mu_k\}_{k \in \mathbb{N}^+}$ is a sequence of times at which the trajectory crosses \mathcal{S} . The mean IPI is defined as

$$\bar{T}_\epsilon \triangleq \mathbb{E}[\Delta_k] \quad (9.35)$$

and the second moment of the IPI is defined as

$$S_\epsilon \triangleq \mathbb{E}[\Delta_k^2]. \quad (9.36)$$

The variance of the IPI is defined as

$$\sigma_\phi^2 \triangleq \mathbb{E}[(\bar{T}_\epsilon - \Delta_k)^2]. \quad (9.37)$$

The moments (9.35)-(9.37) are evaluated under the stationary probability distribution.

It follows immediately that for a given noise level ϵ , we have $\sigma_\phi^2 = S_\epsilon - \Delta_\epsilon^2$.

Remark 5. *Each iso-phase crossing time, $\{\mu_k\}_{k \in \mathbb{N}^+}$, in Definition 6, is a stopping time.*

Remark 6. *Because (9.28) is a diffusion with continuous sample paths, it is possible that when $\epsilon > 0$ a stochastic trajectory $\mathbf{X}(t)$ may make multiple crossings of an iso-phase*

section \mathcal{S} in quick succession. Should this be the case, we condition the crossing times μ_k on completion of successive circuits around the hole in our cylindrical domain. That is, given μ_k , we take μ_{k+1} to be the first return time to \mathcal{S} after having completed at least one half a rotation around the domain.

Chapter 10

Noise Decomposition of the 14-D Stochastic HH Model

Ermentrout et al. [31] studied the variance of the infinitesimal phase response curve for a neuronal oscillator driven by a white noise current, using a four-dimensional version of the Hodgkin-Huxley model as an example. As a corollary result, they obtained an expression for the variance of the interspike interval, by setting the size of the perturbing voltage pulse to zero.

Stochastic shielding [102] allows one to resolve the molecular contributions (per directed edge in the ion channel state transition graph \mathcal{E} , cf. Fig. 3.1) to the variance of ion channel currents [104, 105], and provides a numerical method for accurate, efficient simulation of Langevin models using a small subset of the independent noise forcing (only for the “most important edges”) [95].

Here we combine the stochastic shielding method with Cao et al.’s mean–return-time phase analysis [75] to obtain an analytical decomposition of the molecular sources of timing

variability under current clamp.

Prior analysis of stochastic shielding ([104, 105]) assumed voltage clamp conditions, under which the ion channel state process is a stationary Markov process. Under current clamp, however, fluctuations of channel state determine fluctuations in current, which in turn dictate voltage changes, which then influence channel state transition probabilities, forming a closed loop of statistical interdependence. Therefore, the variance of ISI under current clamp becomes more difficult to analyze. Nevertheless, in this section, we seek a decomposition of the interspike-interval variance into a sum of contributions from each edge $k \in \mathcal{E}$, e.g.

$$\sigma_{\text{ISI}}^2(\epsilon) = \epsilon \sum_{k \in \mathcal{E}} \sigma_{\text{ISI},k}^2 + O(\epsilon^2) \quad (10.1)$$

$$\sigma_{\phi}^2(\epsilon) = \epsilon \sum_{k \in \mathcal{E}} \sigma_{\phi,k}^2 + O(\epsilon^2) \quad (10.2)$$

to leading order as $\epsilon \rightarrow 0$.

Theorem 4 below gives the detailed form of the decomposition. As preliminary evidence for its plausibility, Fig. 10.1 shows the coefficient of variation (standard deviation divided by mean) of the ISI under the stochastic shielding approximation for Langevin model in different scenarios: including noise along a single directed edge at a time (blue bars), or on edges numbered 1 to k inclusive (numbering follows that in Fig. 3.1). For large noise (Fig. 10.1a,c), the effects of noise from different edges combine subadditively. For small noise (Fig. 10.1b,d) contributions of distinct edges to firing variability combine additively. Edges with small contribution to steady-state occupancy under voltage clamp (edges 1-6 for K^+ , edges 1-18 for Na^+ , cf. Fig. 3.1) contribute additively even in the large-noise regime. Thus even in the large-noise regime, stochastic shielding allows accurate simulation of ISI variability using significantly fewer edges for both the sodium and potas-

sium channels.

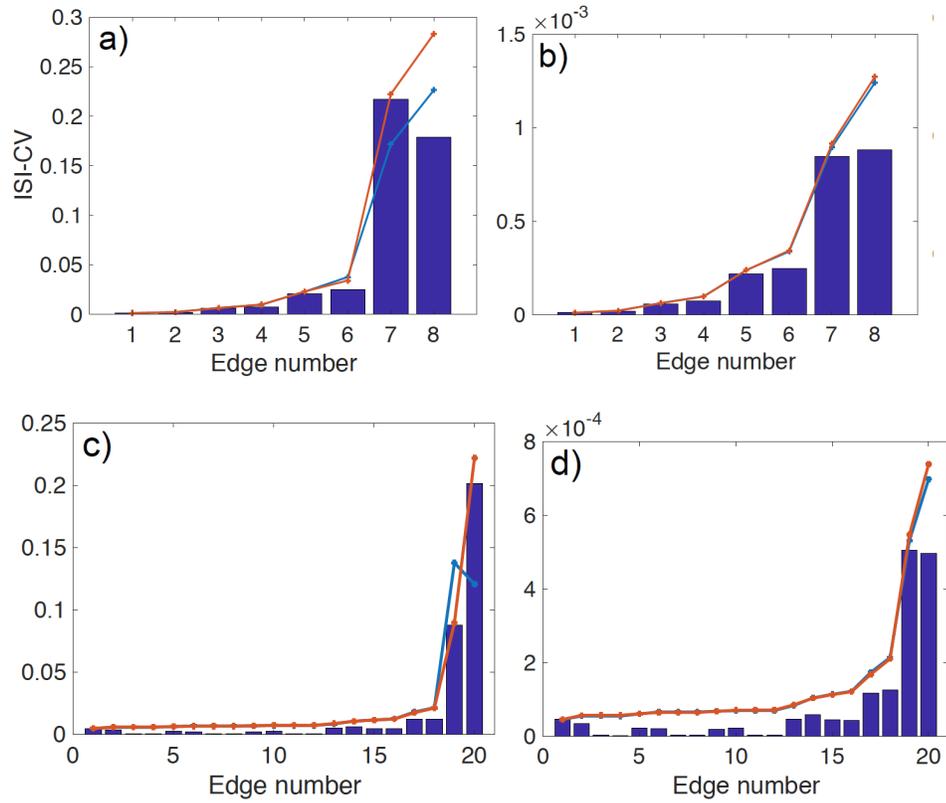


Figure 10.1: Approximate decomposition of interspike interval (ISI) variance into a sum of contributions from each edge for Hodgkin-Huxley model with stochastic K^+ and deterministic Na^+ gates (a,b) or stochastic Na^+ and deterministic K^+ gates (c,d). Bar n shows ISI coefficient of variation (CV) when noise on edge n is included (a,c: $\epsilon = 1$, large noise; b,d: $\epsilon = 0.01$, small noise). Blue line shows the CV of ISI when noise on all edges numbered $\leq n$ are included. Red line shows CV predicted by a linear superposition of the form (10.8).

10.1 Assumptions for Decomposition of the Full Noise Model

Consider a Langevin model for a single-compartment conductance-based neuron (9.28).

We organize the state vector into the voltage component followed by fractional gating

variables as follows:

$$\mathbf{x} = (v, q_1, q_2, \dots, q_N)^\top. \quad (10.3)$$

Here, N is the number of nodes in the union of the ion channel state graphs. For example, for the HH system, $N = 13$, and we would write $q_1 = m_{00}, \dots, q_8 = m_{31}$ for the sodium gating variables, and $q_9 = n_0, \dots, q_{13} = n_4$ for the potassium gating variables. Similarly, we enumerate the K edges occurring in the union of the ion channel state graphs, and write the stoichiometry vector $\zeta_k \in \mathbb{R}^{N+1}$ for transition k , taking source $i(k)$ to destination $j(k)$, in terms of $(N + 1)$ -dimensional unit vectors $\mathbf{e}_i^{N+1} \in \mathbb{R}^{N+1}$ as $\zeta_k = -\mathbf{e}_{i(k)}^{N+1} + \mathbf{e}_{j(k)}^{N+1}$. In order to study the contributions of individual molecular transitions to spike-time variability, we develop asymptotic expansions of the first and second moments of the distribution of iso-phase surface crossing times (iso-phase interval distribution see Def. 6 above) in the small ϵ limit.

Before formally stating the theorem, we make the following assumptions concerning the system (9.28):

A1 We assume the deterministic dynamical system $\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x})$ has an asymptotically, linearly stable limit cycle $\mathbf{x} = \gamma(t)$ with finite period T_0 , and asymptotic phase function $\phi(\mathbf{x})$ defined throughout the limit cycle's basin of attraction such that $\frac{d\phi(\mathbf{x}(t))}{dt} = \frac{2\pi}{T_0}$ along deterministic trajectories, and a well defined infinitesimal phase response curve (iPRC), $\mathbf{Z}(t) = \nabla\phi(\gamma(t))$.

A2 We assume that the $(N + 1) \times K$ matrix \mathcal{G} has the form

$$\mathcal{G}(\mathbf{x}) = \sum_{k=1}^K (\zeta_k \mathbf{r}_k) \sqrt{\alpha_k(v) q_{i(k)}} \quad (10.4)$$

where $\mathbf{r}_k = (\mathbf{e}_k^K)^\top$ is an K -dimensional unit row vector with all zero components

except in the k th entry, $\alpha_k(v)$ is the voltage-dependent per capita transition rate along the k th directed edge, and the $q_{i(t)}$ denote channel state occupancy probabilities as described above (cf. (10.3)).

Remark 7. *The product $\zeta_k \mathbf{r}_k$ is a $(N + 1) \times K$ sparse matrix, containing zeros everywhere except in the k th column. Each column conveys the impact of an independent noise source on the state vector [95].*

A3 We assume that for sufficiently small noise, $0 < |\epsilon| \ll 1$, we have a well defined joint stationary probability distribution in the voltage V and the gating variables Q_1, \dots, Q_N with a well defined mean period \bar{T}_ϵ and mean–return–time phase function $T_\epsilon(\mathbf{x})$. Moreover, we assume that the mean period, the MRT function, and the second moment function all have well defined series expansions:

$$\bar{T}_\epsilon = \bar{T}_0 + \epsilon \bar{T}_1 + O(\epsilon^2) \quad (10.5)$$

$$T_\epsilon(\mathbf{x}) = T_0(\mathbf{x}) + \epsilon T_1(\mathbf{x}) + O(\epsilon^2) \quad (10.6)$$

$$S_\epsilon(\mathbf{x}) = S_0(\mathbf{x}) + \epsilon S_1(\mathbf{x}) + O(\epsilon^2), \quad (10.7)$$

as $\epsilon \rightarrow 0$.

Remark 8. *Note that the expansions (10.5)-(10.7) may break down in the small- ϵ limit for noise-dependent oscillators, such as the heteroclinic oscillator [121] or ecological quasi-cycles [80], but should remain valid for finite-period limit cycles such as the Hodgkin-Huxley system in the periodic spiking regime.*

10.2 Noise Decomposition Theorem

Theorem 4 (Noise Decomposition Theorem). *Let $\mathbf{x} = (V_0, \mathbf{Q}_0) \in \mathcal{S}_0$ be the point on the deterministic limit cycle such that $\phi(\mathbf{x}) = 0$ (i.e. assigned to “phase zero”), and let $\mathbb{E}^{\mathbf{x}}$ denote expectation with respect to the law of trajectories with initial condition \mathbf{x} , for fixed $\epsilon \geq 0$. Under assumptions A1-A3, the variance σ_ϕ^2 of the \mathcal{S}_0 -isochron crossing times (iso-phase intervals, or IPI) for conductance-based Langevin models (eqn. (6.42)) decomposes into additive contributions from each channel-state transition, in the sense that*

$$\sigma_\phi^2 = \sum_{k \in \text{all edges}} \sigma_{\phi,k}^2 \quad (10.8)$$

$$= \epsilon \sum_k \int_0^{\bar{T}_0} \mathbb{E}^{\mathbf{x}} \left(\alpha_k(V(t)) Q_{i(k)}(t) \left(\zeta_k^\top \tilde{\mathbf{Z}}(\mathbf{X}(t)) \right)^2 \right) dt + O(\epsilon^2), \quad (10.9)$$

as $\epsilon \rightarrow 0^+$. The function $\mathbf{X}(t) = (V(t), Q_1(t), \dots, Q_N(t))^\top$ denotes a stochastic trajectory of (9.28) with initial condition \mathbf{x} .

Remark 9. *The theorem holds independently of the choice of the initial point \mathbf{x} on the deterministic limit cycle, in the sense that choosing a different base point would just shift the endpoints of the interval of integration; since the deterministic limit cycle is periodic with period \bar{T}_0 , the resulting expression for σ_ϕ^2 is the same. See Corollary 7.*

Remark 10. *The proof relies on Dynkin’s formula, first–passage-time moment calculations, and a small noise expansion. The right hand side of (10.8) leads to an approximation method based on sampling stochastic limit cycle trajectories, which we show below gives an accurate estimate for σ_ϕ^2 .*

Remark 11. *Although the interspike intervals (ISI) determined by voltage crossings are not strictly identical to the iso-phase intervals (IPI) defined by level crossings of the function*

$T_\epsilon(\mathbf{x})$, we nevertheless expect that the variance of the IPI, and their decomposition, will provide an accurate approximation to the variance of the ISI. In §11 we show numerically that the decomposition given by (10.8) predicts the contribution of different directed edges to the voltage-based ISIs with a high degree of accuracy.

Before proving the theorem, we state and prove two ancillary lemmas.

Lemma 5. Fix a cylindrical domain \mathcal{D}_0 (as in equation (9.33)) and an iso-phase section \mathcal{S}_0 transverse to the vector field \mathbf{F} . If the mean period \bar{T}_ϵ and MRT function $T_\epsilon(\mathbf{x})$ have Taylor expansions (10.5) and (10.6), then the unperturbed isochron function T_0 and the sensitivity of the isochron function to small noise T_1 satisfy

$$\mathbf{F}(\mathbf{x}) \cdot \nabla T_0(\mathbf{x}) = -1, \quad (10.10)$$

$$\mathbf{F}(\mathbf{x}) \cdot \nabla T_1(\mathbf{x}) = -\frac{1}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0(\mathbf{x}), \quad (10.11)$$

$$\sum_{ij} n_i (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_j T_1 \Big|_{\partial\mathcal{D}} = 0, \quad (10.12)$$

where $T_1 \rightarrow T_1 + \bar{T}_1$ and $T_0 \rightarrow T_0 + \bar{T}_0$ across \mathcal{S}_0 , and \mathbf{n} is the outward normal to the boundary $\partial\mathcal{D}$.

Note that \bar{T}_1 may be determined from the stationary solution of the forward equation for $0 < \epsilon$, or through Monte Carlo simulations (in some cases $\bar{T}_1 \equiv 0$).

Proof. For all noise levels $\epsilon \geq 0$, from [43] (Chapter 5, equation 5.5.19), the MRT function $T_\epsilon(x)$ satisfies

$$\mathcal{L}^\dagger [T_\epsilon] = \mathbf{F} \cdot \nabla T_\epsilon + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_\epsilon = -1, \quad (10.13)$$

together with adjoint reflecting boundary conditions at the edges of the domain \mathcal{D} with

outward normal vector \mathbf{n}

$$\sum_{ij} n_i (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_j T_\epsilon \Big|_{\partial\mathcal{D}} = 0 \quad (10.14)$$

and the jump condition is specified as follows. When x increases across the reference section \mathcal{S} in the “forward direction”, i.e., in a direction consistent with the mean flow in forwards time, the function $T_\epsilon \rightarrow T_\epsilon + \bar{T}_\epsilon$. Note that since $T_0 \rightarrow T_0 + \bar{T}_0$, we also have $T_1 \rightarrow T_1 + \bar{T}_1$ across the same Poincaré section, for consistency.

Substituting the expansion (10.6) into (10.13) gives

$$\begin{aligned} -1 = & \mathbf{F} \cdot \nabla (T_0(\mathbf{x}) + \epsilon T_1(\mathbf{x}) + O(\epsilon^2)) + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 (T_0(\mathbf{x}) + \epsilon T_1(\mathbf{x}) + O(\epsilon^2)), \\ & (10.15) \end{aligned}$$

$$\begin{aligned} = & \mathbf{F} \cdot \nabla T_0(\mathbf{x}) + \mathbf{F} \cdot \nabla (\epsilon T_1(\mathbf{x}) + O(\epsilon^2)) + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0(\mathbf{x}) \\ & + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 (\epsilon T_1(\mathbf{x}) + O(\epsilon^2)) \end{aligned} \quad (10.16)$$

Note that, when $\epsilon = 0$

$$\mathbf{F} \cdot \nabla T_0(\mathbf{x}) = -1, \quad (10.17)$$

consistent with T_0 being equal to minus the asymptotic phase of the limit cycle (up to an additive constant). On the other hand, for $\epsilon \neq 0$, by comparing the first order term, the sensitivity of the isochron function to small noise T_1 satisfies

$$\mathbf{F} \cdot \nabla T_1(\mathbf{x}) = -\frac{1}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0, \quad \sum_{ij} n_i (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_j T_1 \Big|_{\partial\mathcal{D}} = 0, \quad (10.18)$$

where $T_1 \rightarrow T_1 + \bar{T}_1$ across \mathcal{S} , and \mathbf{n} is the outward normal to the boundary $\partial\mathcal{D}$, thus we

proved Lemma 5. □

Our next lemma concerns the second moment of the first passage time from a point $\mathbf{x} \in \mathcal{D}$ to a given iso-phase section \mathcal{S}_0 , that is, $S_\epsilon(\mathbf{x}) \triangleq E[\tau(\mathbf{x}, \mathcal{S}_0)^2]$, cf. (9.17).

Lemma 6. *Suppose the assumptions of Lemma 5 hold, and assume in addition that S_ϵ has a Taylor expansion (10.7) for small ϵ . The second moment, $S_0(\mathbf{x})$, of the first passage time $\tau(\mathbf{X}, \mathcal{S})$ from a point \mathbf{x} to a given isochron section $\mathcal{S}_0 = \{T_\epsilon(\mathbf{x}) = \text{const}\}$, and its first order perturbation, $S_1(\mathbf{x})$, satisfy*

$$\mathbf{F} \cdot \nabla S_0 = -2T_0 \tag{10.19}$$

$$\mathbf{F} \cdot \nabla S_1 + \frac{1}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 S_0 = -2T_1. \tag{10.20}$$

Proof. Following [43] (Chapter 5, equation 5.5.19), the second moment $S_\epsilon(\mathbf{x})$ of the first passage time from a point \mathbf{x} to a given isochron $T_\epsilon(\mathbf{x}) = \text{const}$, satisfies

$$\mathcal{L}^\dagger[S_\epsilon] := \mathbf{F} \cdot \nabla S_\epsilon + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 S_\epsilon = -2T_\epsilon.$$

Substituting in the Taylor expansions (10.5)-(10.7), we have to order $O(\epsilon)$

$$\mathbf{F} \cdot \nabla(S_0 + \epsilon S_1) + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2(S_0 + \epsilon S_1) = -2(T_0 + \epsilon T_1) + O(\epsilon^2). \tag{10.21}$$

Setting $\epsilon = 0$, we see that

$$\mathbf{F} \cdot \nabla S_0 = -2T_0. \tag{10.22}$$

For $\epsilon > 0$, the first order terms yield

$$\mathbf{F} \cdot \nabla S_1 + \frac{1}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 S_0 = -2T_1. \quad (10.23)$$

Therefore, we complete the proof of Lemma 6. \square

10.3 Proof of Theorem 4

Proof. We divide the proof of the Theorem into three steps.

1. First, we will calculate the infinitesimal generator for the variance of the iso-phase interval (IPI).

For fixed noise level $\epsilon > 0$, the variance of IPI, σ_ϕ^2 is equal to the expected value of $V_\epsilon = S_\epsilon - T_\epsilon^2$, evaluated at the isochron $T = \text{const} + \bar{T}$. Note that when $\epsilon = 0$, the system is deterministic and the iso-phase interval has a zero variance, i.e., $V_0 \equiv 0$. Expanding $S_\epsilon = S_0 + \epsilon S_1 + O(\epsilon^2)$ and $V_\epsilon = \epsilon V_1 + O(\epsilon^2)$ to first order in $\epsilon \ll 1$, then

$$V_\epsilon = V_0 + \epsilon V_1 + O(\epsilon^2) \quad (10.24)$$

$$= S_\epsilon - T_\epsilon^2 \quad (10.25)$$

$$= S_0 + \epsilon S_1 + O(\epsilon^2) - (T_0(\mathbf{x}) + \epsilon T_1(\mathbf{x}) + O(\epsilon^2))^2 \quad (10.26)$$

$$= S_0 - T_0^2 + \epsilon(S_1 - 2T_0 T_1) + O(\epsilon^2), \quad (10.27)$$

thus,

$$S_0 = T_0^2 \quad (10.28)$$

$$S_1 = V_1 + 2T_0 T_1. \quad (10.29)$$

Plug the above results into equation (10.23) (Lemma 6), we can obtain that

$$\mathbf{F} \cdot \nabla(V_1 + 2T_0T_1) + \frac{1}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0^2 = -2T_1. \quad (10.30)$$

By the product rule and use equations (10.17), and (10.18),

$$\mathbf{F} \cdot \nabla(2T_0T_1) = 2T_1\mathbf{F} \cdot \nabla(T_0) + 2T_0\mathbf{F} \cdot \nabla(T_1) \quad (10.31)$$

$$= -2T_1 - T_0 \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0. \quad (10.32)$$

Therefore,

$$\mathbf{F} \cdot \nabla V_1 - 2T_1 - T_0 \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0 + \frac{1}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0^2 = -2T_1. \quad (10.33)$$

Since $\partial_{ij}^2 T_0^2 = \partial_i(2T_0\partial_j T_0) = 2\partial_i T_0\partial_j T_0 + 2T_0\partial_{ij}^2 T_0$, it follows that

$$\begin{aligned} \mathbf{F} \cdot \nabla V_1 - T_0 \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 T_0 + \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_i T_0 \partial_j T_0 + \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} T_0 \partial_{ij}^2 T_0 &= 0 \\ \mathbf{F} \cdot \nabla V_1 &= - \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_i T_0 \partial_j T_0 \end{aligned} \quad (10.34)$$

Finally,

$$\mathcal{L}^\dagger[V_\epsilon] = \mathcal{L}^\dagger[V_0 + \epsilon V_1 + O(\epsilon^2)] \quad (10.35)$$

$$= \epsilon \mathcal{L}^\dagger[V_1] + O(\epsilon^2) \quad (10.36)$$

$$= \epsilon \left(\mathbf{F} \cdot \nabla V_1 + \frac{\epsilon}{2} \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_{ij}^2 V_1 \right) + O(\epsilon^2) \quad (10.37)$$

$$= \epsilon (\mathbf{F} \cdot \nabla V_1) + O(\epsilon^2) \quad (10.38)$$

$$= -\epsilon \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_i T_0 \partial_j T_0 + O(\epsilon^2), \quad (10.39)$$

where we used $V_0 \equiv 0$ and applied equation (10.34).

2. Secondly, we will show that for first-order transition networks underlying the molecular ion channel process, the decomposition $\mathcal{G}\mathcal{G}^\top = \sum_{k \in \mathcal{E}} \mathcal{G}_k \mathcal{G}_k^\top$ is exact.

To see this, note that \mathcal{G} can be written as a sum of 29 sparse matrix with one zero matrix and 28 rank one matrix. The k th rank one matrix consists of the transition due to the k th edge and there are 28 edges in the 14-D HH model. The k th column of the k th rank one matrix equals to a stoichiometry vector times the square root of the corresponding state occupancy and zeros otherwise. For example, the k th column of \mathcal{G} is given by

$$\mathcal{G}_k = \zeta_k \sqrt{\alpha_k(v) \mathbf{X}_{i(k)}},$$

where ζ_k is the stoichiometry vector, α_k is the voltage-dependent *per capita* transition rate, and $\mathbf{X}_{i(k)}$ is the population vector component at the source node $i(k)$ for transition number k .

$$\mathcal{G}\mathcal{G}^\top = (\mathcal{G}_1 + \mathcal{G}_2 + \cdots + \mathcal{G}_{28})(\mathcal{G}_1 + \mathcal{G}_2 + \cdots + \mathcal{G}_{28})^\top \quad (10.40)$$

$$= \sum_{k=1}^{28} \mathcal{G}_k \mathcal{G}_k^\top \quad (10.41)$$

$$= \sum_{k=1}^{28} \alpha_k(v) \mathbf{X}_{i(k)} \zeta_k \zeta_k^\top \quad (10.42)$$

where (10.41) holds because $\mathcal{G}_i \mathcal{G}_j^\top = 0$ when $i \neq j$.

Note that $\partial_i T_0 \partial_j T_0 = \omega^{-2} \partial_i \phi(\mathbf{x}) \partial_j \phi(\mathbf{x}) = \tilde{\mathbf{Z}}_i(\mathbf{x}) \tilde{\mathbf{Z}}_j(\mathbf{x})$, with $\omega \equiv 2\pi/\bar{T}_0$, because ϕ

is normalized to range from 0 to 2π , and T_0 ranges from 0 to \bar{T}_0 .

$$\sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_i T_0 \partial_j T_0 = \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \tilde{\mathbf{Z}}_i \tilde{\mathbf{Z}}_j \quad (10.43)$$

$$= \sum_{k=2}^{29} \sum_{ij} (\mathcal{G}_k \mathcal{G}_k^\top)_{ij} \tilde{\mathbf{Z}}_i \tilde{\mathbf{Z}}_j \quad (10.44)$$

$$= \sum_{k=1}^{28} \left(\alpha_k(v) \mathbf{X}_{i(k)} \sum_{ij} (\zeta_k \zeta_k^\top)_{ij} \tilde{\mathbf{Z}}_i \tilde{\mathbf{Z}}_j \right) \quad (10.45)$$

$$= \sum_{k=1}^{28} \alpha_k(v) \mathbf{X}_{i(k)} \left[\tilde{\mathbf{Z}}_{i(k)}^2 + \tilde{\mathbf{Z}}_{j(k)}^2 - 2\tilde{\mathbf{Z}}_{i(k)} \tilde{\mathbf{Z}}_{j(k)} \right] \quad (10.46)$$

$$= \sum_{k=1}^{28} \alpha_k(v) \mathbf{X}_{i(k)} \left[\tilde{\mathbf{Z}}_{i(k)} - \tilde{\mathbf{Z}}_{j(k)} \right]^2 \quad (10.47)$$

$$= \sum_{k=1}^{28} \alpha_k(v) \mathbf{X}_{i(k)} \left(\zeta_k^\top \tilde{\mathbf{Z}} \right)^2, \quad (10.48)$$

where $i(k)$ and $j(k)$ are the source and sink nodes for transition number k . Equation (10.46) holds because the k th edge only involves two nodes.

3. Finally, we will apply Dynkin's formula to complete the rest of the proof.

For a stopping time $\tau(\mathbf{x})$ with $\mathbb{E}^{\mathbf{x}}(\tau) < \infty$, by Dynkin's formula (9.31), the expected IPI variance starting from \mathbf{x} is

$$\mathbb{E}^{\mathbf{x}}(V_\epsilon(\mathbf{X}(\tau))) = V_\epsilon(\mathbf{x}) + \mathbb{E}^{\mathbf{x}} \left(\int_0^\tau \mathcal{L}^\dagger[V_\epsilon(\mathbf{X}(s))] ds \right) \quad (10.49)$$

The first return time τ is the time at which the trajectory $\mathbf{X}(t)$ first returns to the isochron \mathcal{S}_0 , therefore $\mathbf{X}(\tau) \in \mathcal{S}_0$ and the time left to reach \mathcal{S}_0 from the random location $\mathbf{X}(\tau)$ is guaranteed to be zero. That is, $V_\epsilon(\mathbf{X}(\tau)) = 0$ with probability 1. Hence, $\mathbb{E}^{\mathbf{x}}(V_\epsilon(\mathbf{X}(\tau))) \equiv 0$ for all $\mathbf{x} \in \mathcal{S}_0$.

Fix a mean–return-time isochron \mathcal{S}_0 , the mean return time from any initial location $\mathbf{x} \in \mathcal{S}_0$ back to \mathcal{S}_0 , after completing one rotation is exactly \bar{T}_ϵ , by construction. However, in principle, the *variance* of the return time might depend on the initial location within the isochron. We next show that, to leading order in ϵ , this is not the case, that is, the MRT isochrons have uniform first *and* second moment properties.

Using equations (10.39), (10.48) and (10.49), we obtain

$$V_\epsilon(\mathbf{x}) = -\mathbb{E}^{\mathbf{x}} \left(\int_0^\tau \mathcal{L}^\dagger [V_\epsilon(\mathbf{X}(s))] ds \right) \quad (10.50)$$

$$= \mathbb{E}^{\mathbf{x}} \left(\int_0^\tau \epsilon \sum_{ij} (\mathcal{G}\mathcal{G}^\top)_{ij} \partial_i T_0 \partial_j T_0 ds \right) + \mathcal{O}(\epsilon^2) \quad (10.51)$$

$$= \epsilon \sum_{k=1}^{28} \mathbb{E}^{\mathbf{x}} \left(\int_0^\tau \alpha_k(v) \mathbf{X}_{i(k)} \left(\zeta_k^\top \tilde{\mathbf{Z}} \right)^2 ds \right) + \mathcal{O}(\epsilon^2), \quad (10.52)$$

where the integrals are evaluated along a stochastic trajectory $\mathbf{X}(t)$ with $\mathbf{X}(0) = \mathbf{x}$ and $\mathbf{X}(\tau) \in \mathcal{S}_0$, one rotation later. Holding the deterministic zero-phase isochron \mathcal{S}_0 fixed, and choosing an arbitrary point $\mathbf{y} \in \mathcal{D}$, we have, by definition,

$$\mathbb{E}^{\mathbf{y}}[\tau(\mathbf{y})] = T_\epsilon(\mathbf{y}) = T_0(\mathbf{y}) + \epsilon T_1(\mathbf{y}) + \mathcal{O}(\epsilon^2). \quad (10.53)$$

Therefore, starting from an initial condition $\mathbf{x} \in \mathcal{S}_0$ *one period earlier*, we have

$$\mathbb{E}^{\mathbf{x}} \left(\int_0^\tau \alpha_k(v) \mathbf{X}_{i(k)} \left(\zeta_k^\top \tilde{\mathbf{Z}} \right)^2 ds \right) = \mathbb{E}^{\mathbf{x}} \left(\int_0^{\bar{T}_0} \alpha_k(v) \mathbf{X}_{i(k)} \left(\zeta_k^\top \tilde{\mathbf{Z}} \right)^2 ds \right) + \mathcal{O}(\epsilon^2). \quad (10.54)$$

This relation follows immediately from our assumptions, because, for $\mathbf{x} \in \mathcal{S}_0$,

$$\left| \mathbb{E}^{\mathbf{x}} \left(\int_0^\tau \alpha_k(V(s)) \mathbf{X}_{i(k)}(s) \left(\zeta_k^\top \tilde{\mathbf{Z}}(s) \right)^2 ds \right) - \mathbb{E}^{\mathbf{x}} \left(\int_0^{\bar{T}_0} \alpha_k(V(s)) \mathbf{X}_{i(k)}(s) \left(\zeta_k^\top \tilde{\mathbf{Z}}(s) \right)^2 ds \right) \right|$$

$$= \left| \mathbb{E}^{\mathbf{x}} \left(\int_{\bar{T}_0}^\tau \alpha_k(V(s)) \mathbf{X}_{i(k)}(s) \left(\zeta_k^\top \tilde{\mathbf{Z}}(s) \right)^2 ds \right) \right| \quad (10.55)$$

$$\leq C_1 |\mathbb{E}^{\mathbf{x}}(\tau(\mathbf{x}) - \bar{T}_0)| = C_1 |\mathbb{E}^{\mathbf{x}}(\tau(\mathbf{x})) - \mathbb{E}^{\mathbf{x}}(\bar{T}_0)| = C_1 |\bar{T}_\epsilon - \bar{T}_0| \quad (10.56)$$

$$= \epsilon C_1 \bar{T}_1 + \mathcal{O}(\epsilon^2). \quad (10.57)$$

Here C_1 is a positive constant bounding the integrand $\alpha_k(v(t)) \mathbf{X}_{i(k)}(t) \left(\eta_k^\top \tilde{\mathbf{Z}}(t) \right)^2$. From Remark 3, $\alpha_k \leq \alpha_{\max}$. By definition, $0 \leq \mathbf{X}_i \leq 1$ for each i . For each edge k , $|\zeta_k| = \sqrt{2}$. Since $\tilde{\mathbf{Z}}$ is continuous and periodic, $|\tilde{\mathbf{Z}}|$ is bounded by some constant z_{\max} . Therefore setting $C_1 = \sqrt{2} \alpha_{\max} z_{\max}$ satisfies (10.57).

Because the initial point $\mathbf{x} \in \mathcal{S}_0$ was located at an arbitrary radius along the specified mean–return-time isochron, the calculation above shows that $\sigma_\phi^2 = \mathbb{E}[V_\epsilon(\mathbf{x}) | \mathbf{x} \in \mathcal{S}_0]$ is uniform across the isochron \mathcal{S}_0 , to first order in ϵ . Thus, for small noise levels, the MRT isochrons enjoy not only a uniform mean return time, but also a uniform variance in the return time, at least in the limit of small noise.

Finally, we note that σ_ϕ^2 (equivalently, and $V_\epsilon(\mathbf{x})$) combine a sum of contributions over a finite number of edges. From equations (10.52) and (10.54), the variance of the inter-phase interval is given by

$$\sigma_\phi^2 = \epsilon \sum_{k=1}^{28} \mathbb{E}^{\mathbf{x}} \left(\int_0^{\bar{T}_0} \alpha_k(V(s)) \mathbf{X}_{i(k)}(s) \left(\zeta_k^\top \tilde{\mathbf{Z}}(s) \right)^2 ds \right) + \mathcal{O}(\epsilon^2). \quad (10.58)$$

To complete the proof, note that (10.9) follows from (10.59) by exchange of expectation $\mathbb{E}^{\mathbf{x}}[\cdot]$ with (deterministic) integration $\int_0^{\bar{T}_0}[\cdot] dt$. This completes the proof of Theorem 4.

□

The choice of the initial reference point \mathbf{x} or isochron \mathcal{S}_0 in (10.59) was arbitrary and the variance of IPI is uniform to the first order. Therefore, the inter-phase-interval variance may be uniform (to first order) almost everywhere in \mathcal{D} . We can then replace the integral around the limit cycle in (10.59) with an integral over \mathcal{D} with respect to the stationary probability distribution. Thus we have the following

Corollary 7. *Under the assumptions of Theorem 4, the inter-phase-interval variance satisfies*

$$\sigma_{\phi}^2 = \epsilon \bar{T}_0 \sum_{k=1}^{28} \mathbb{E} \left(\alpha_k(V) \mathbf{X}_{i(k)} \left(\zeta_k^{\top} \tilde{\mathbf{Z}}(\mathbf{X}) \right)^2 \right) + \mathcal{O}(\epsilon^2), \quad (10.59)$$

as $\epsilon \rightarrow 0$, where \mathbb{E} denotes expectation with respect to the stationary probability density for (9.28).

Remark 12. *Because the variance of the IPI, σ_{ϕ}^2 , is uniform regardless the choice of the reference iso-phase section, we will henceforth refer it as σ_{IPI}^2 throughout the rest of this thesis.*

Now we can say we have generalized the edge important measure introduced in [105] for the voltage-clamp case to the current clamp case with weak noise. In the next chapter we leverage Theorem 4 to estimate the inter-phase interval variance in two different ways: by averaging over one period of the deterministic limit cycle (compare (10.58)) or by averaging over a long stochastic simulation (compare (10.59)). As we will see below,

both methods give excellent agreement with direct measurement of the inter-phase interval variance.

Chapter 11

Decomposition of the Variance of Interspike Intervals

Theorem 4 and Corollary 7 assert that for sufficiently weak levels of channel noise, the contributions to inter-phase interval variance made by each individual edge in the channel state transition graph (cf. Fig. 3.1) combine additively. Moreover, the relative sizes of these contributions provide a basis for selecting a subset of noise terms to include for running efficient yet accurate Langevin simulations, using the stochastic shielding approximation [95]. In this chapter, we test and illustrate several aspects of these results numerically.

First, we confront the fact that the inter-*phase*-intervals and the inter-*spike*-intervals are not equivalent, since iso-voltage surfaces do not generally coincide with isochronal surfaces [128]. Indeed, upon close examination of the ISI variance in both real and simulated nerve cells, we find that the voltage-based σ_{ISI}^2 is not constant, as a function of voltage, while the phase-based σ_{IPi}^2 remains the same regardless of the choice of reference isochron. Nevertheless, we show that the voltage-based ISI variance is well approximated – to within

a few percent – by the phase-based IPI variance, and therefore, the linear decomposition of Theorem 4 approximately extends to the ISI variance as well.

Second, after showing that the linear decomposition of the ISI variance holds at sufficiently small noise levels, we explore the range of noise levels over which the linear superposition of edge-specific contributions to ISI variance holds. Consistent with the basic stochastic shielding phenomenon, we find that the variability resulting from noise along edges located further from the observable transitions scales linearly with noise intensity, ϵ even for moderate noise levels, while the linear scaling of eqn. (10.58) breaks down sooner with increasing noise for edges closer to observable transitions.

Finally, we explore the accuracy of a reduced representation using only the two most important edges from the K^+ channel and the four most important edges from the Na^+ channel, over a wide range of noise intensities. Here, we find that removing the noise from all but these six edges still gives an accurate representation of the ISI variance far beyond the weak noise regime, despite the apparent breakdown of linearity.

In this section, the variance of ISIs and IPIs are calculated to compare with the predictions using Theorem 4. First, we numerically show that there is a small-noise region within which Theorem 4 holds, for each individual edge, as well as for the whole Langevin model (cf. (9.28)). We have two numerical approaches to evaluating the theoretical contributions. The first method involves integrating once around the deterministic limit cycle while evaluating the local contribution to timing variance at each point along the orbit. This approach derives from the theorem, cf. (10.9) or (10.58), which we refer as the “limit cycle prediction”. The second approach derives from the corollary, (10.59): we average the expected local contribution to timing variation over a long stochastic trajectory. More specifically, equation (10.59) gives a theoretical value of the average leading-order

contribution mass function, \mathcal{P}_k , for the k^{th} edge, as

$$\mathcal{P}_k := \mathbb{E} \left[\alpha_k(V) \mathbf{X}_{i(k)} \left(\zeta_k^\top \tilde{\mathbf{Z}}(\mathbf{X}) \right)^2 \right], \quad (11.1)$$

where $\mathbb{E}(\cdot)$ is the mean with respect to the stationary probability distribution of the stochastic limit cycle. Given a sample trajectory $\mathbf{X}(t)$, we approximate the iPRC near the limit cycle, $\tilde{\mathbf{Z}}(\mathbf{X}(t))$, by using the phase response curve of the deterministic limit cycle

$$\tilde{\mathbf{Z}}(\mathbf{X}(t)) \approx \hat{\mathbf{Z}}(\mathbf{X}(t)) \triangleq \mathbf{Z} \left(\underset{s}{\operatorname{argmin}} \left| \left(\gamma(s) - \mathbf{X}(t) \right)^\top \mathbf{Z}(s) \right| \right), \quad (11.2)$$

where γ is a point on the deterministic limit cycle and \mathbf{Z} is the infinitesimal phase response curve on the limit cycle (cf. §9.3). The predicted contribution of the k^{th} edge to the IPI variance with average period T_0 , is therefore

From Corollary 7 we have

$$\sigma_{\text{IPI}}^2 \approx \epsilon \bar{T}_0 \sum_k \mathcal{P}_k. \quad (11.3)$$

We call \mathcal{P}_k the *point mass prediction* for the contribution of the k th edge to the inter-phase interval variance.

For small noise, both approaches give good agreement with the directly measured IPI variance, as we will see in Fig. 11.3.

To numerically calculate the contribution for each directed transition in Fig. 3.1, we apply the stochastic shielding (SS) technique proposed by [102], simulating the Langevin process with noise from all but a single edge suppressed. Generally speaking, the SS method approximates the Markov process using fluctuations from only a subset of the transitions, often the observable transitions associated to the opening states. Details about how stochastic shielding can be applied to the $14 \times 28\text{D}$ Langevin model is discussed in

§6.5.

All numerical simulations for the Langevin models use the same set of parameters, which are specified in Tab. B.1 with given noise level ϵ in eqn. (9.28). We calculate the following quantities: the point mass prediction \mathcal{P}_k , using exact stochastic trajectories (11.1); the predicted contributions by substituting the stochastic terms in (10.58) with the deterministic limit cycle; the variance and standard deviation of the interspike intervals (σ_{ISI}^2); and the variance and standard deviation of the isophase intervals (σ_{IPI}^2).

In addition to numerical simulations, we will also present several observations of experimental recordings. Data in Fig. 11.1 and Fig. 11.2 were recorded *in vitro* in Dr. Friel's lab from intact wild type Purkinje cells with synaptic input blocked, with voltage sampled at 20KHz and digitized at 1/32 mV resolution. We analyzed fourteen different voltage traces from cerebellar Purkinje cells from wild type mice, and seventeen from mice with the *leaner* mutation. The average number of full spike oscillations is roughly 1200 for wild type PCs (fourteen cells) and 900 for leaner mutation (seventeen cells).

11.1 Observations on σ_{ISI}^2 , and σ_{IPI}^2

When analyzing voltage recordings from *in vitro* Purkinje cells (PCs) and from simulation of the stochastic HH model, we have the following observations. First, given a particular (stochastic) voltage trace, the number of interspike intervals (cf. Def. 5) varies along with the change in voltage threshold used for identifying spikes. Second, within a range of voltage thresholds for which the number of ISIs is constant, the variance of the interspike interval distribution, σ_{ISI}^2 (cf. Def. 5), which is obtained directly from the voltage recordings, nevertheless varies as a function of the threshold used to define the spike times. Thus the ISI variance, a widely studied quantity in the field of computational neuroscience

[57, 74, 86, 112, 118], is not invariant with respect to the choice of voltage threshold. To our knowledge this observation has not been previously reported in the neuroscience literature.¹

Fig. 11.1 plots the histogram of voltage from a wild type PC and number of spikes corresponding to voltage threshold (V_{th}) in the range of $[-60, -10]$ mV. Setting the threshold excessively low or high obviously will lead to too few (or no) spikes. As the threshold increases from excessively low values, the counts of threshold-crossing increases. For example, when V_{th} is in the after hyper-polarization (AHP) range (roughly $-58 \lesssim V_{th} \lesssim -48$ mV in Fig. 11.1) the voltage trajectory may cross the threshold multiple times before it finally spikes. As illustrated in Fig. 11.1, the number of spikes is not a constant as the threshold varies, therefore, the mean and variance of ISI are not well-defined in the regions where extra spikes are counted. To make the number of spikes accurately reflect the number of full oscillation cycles, in what follows we will only use thresholds in a voltage interval that induces the correct number of spikes. Note that, for a given voltage trace and duration (T_{tot}), if two voltage threshold generate the same number of spikes (N_{spike}), the mean ISI would be almost identical, approximately T_{tot}/N_{spike} . This observation holds for both experimental recordings and numerical simulations.

Next we address the sensitivity of the interspike interval to the voltage threshold, within the range over which the number of ISIs is invariant.

From the earliest days of quantitative neurophysiology, the extraction of spike timing information from voltage traces recorded via microelectrode has relied on setting a fixed voltage threshold (originally called a Schmitt trigger, after the circuit designed by O.H. Schmitt [106]). To our knowledge, it has invariably been assumed that the choice of the

¹Throughout this section, we use the term “threshold” in the data analysis sense of a Schmitt trigger [106], rather than the physiological sense of a spike generation mechanism.

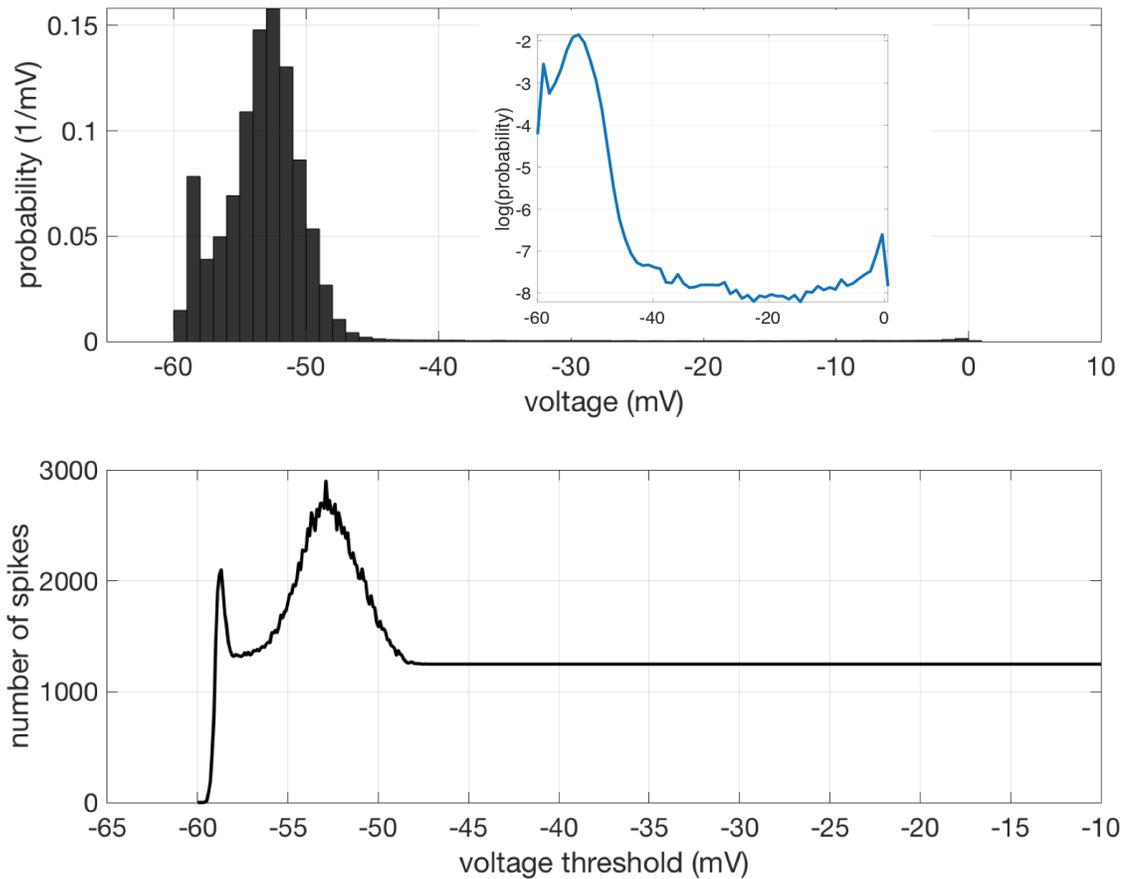


Figure 11.1: Histogram of voltage and number of spikes as a function of voltage threshold (V_{th}) for one wild type Purkinje cell (same data as in Fig. 4.1A & C). The number of ISIs is found by counting up-crossing times as defined in Def. 5. For this particular trajectory, the AHP phase locates roughly in the interval $[-58, -48]$ mV. The trajectory has 1248 full oscillation cycles. When V_{th} is near -60 mV, it captures fewer spikes than the true value, and when $V_{th} \in [-57, -48]$, it tends to overestimate the number of spikes. For $V_{th} \in [-48, -10]$, the number of spikes is a constant (1248) that matches the number of full oscillations.

threshold or trigger level was immaterial, provided it was high enough to avoid background noise and low enough to capture every action potential [44, 84]. This assumption is generally left implicit. Here, we show that, in fact, the choice of the trigger level (the voltage threshold used for identifying spike timing) can cause a change in the variance of the interspike interval for a given spike train by as much as 5%.

Fig. 11.2 provides evidence both from experimental traces recorded *in vitro*, and from numerical simulations, that σ_{ISI}^2 is sensitive to the voltage threshold defining spike times. In Fig. 11.2 A, we superimpose ISI standard deviations from fourteen wild type Purkinje cells, plotted as functions of the the trigger voltage V_{th} . We rescale each plot by the standard deviation of the ISI for each cell at $V_{\text{th}} = -20$ mV, which we define as $\bar{\sigma}$. As shown in Fig. 11.2 A, the cells recorded *in vitro* have a clear variability in the standard deviation as the voltage threshold changes. Specifically, the standard deviation of ISI gradually increases as voltage threshold increases and then remains constant as the threshold approaches the peak of the spikes. Two of the cells have larger variations in the standard deviation, with roughly a 3% – 4% change; nine of them have a 1% – 3% change; and three of them show 0.1% – 1% change.

We applied a similar analysis to seventeen PCs with leaner mutations. In this case, one cell had a variation of roughly 1% in the standard deviation, five cells with variations around 0.2%, and the remaining without an obvious change (data not shown). This difference between cells derived from wild type and leaner mutant mice is an interesting topic for future study.

We observe a similar variability of σ_{ISI}^2 in numerical simulations using our stochastic Langevin HH model (cf. eqn. (9.28)). Fig. 11.2 B and C plots two examples showing the change in σ_{ISI}^2 as voltage threshold varies. For a given noise level (ϵ) and a voltage threshold (V_{th}), a single run simulates a total time of 9000 milliseconds (ms), with a time

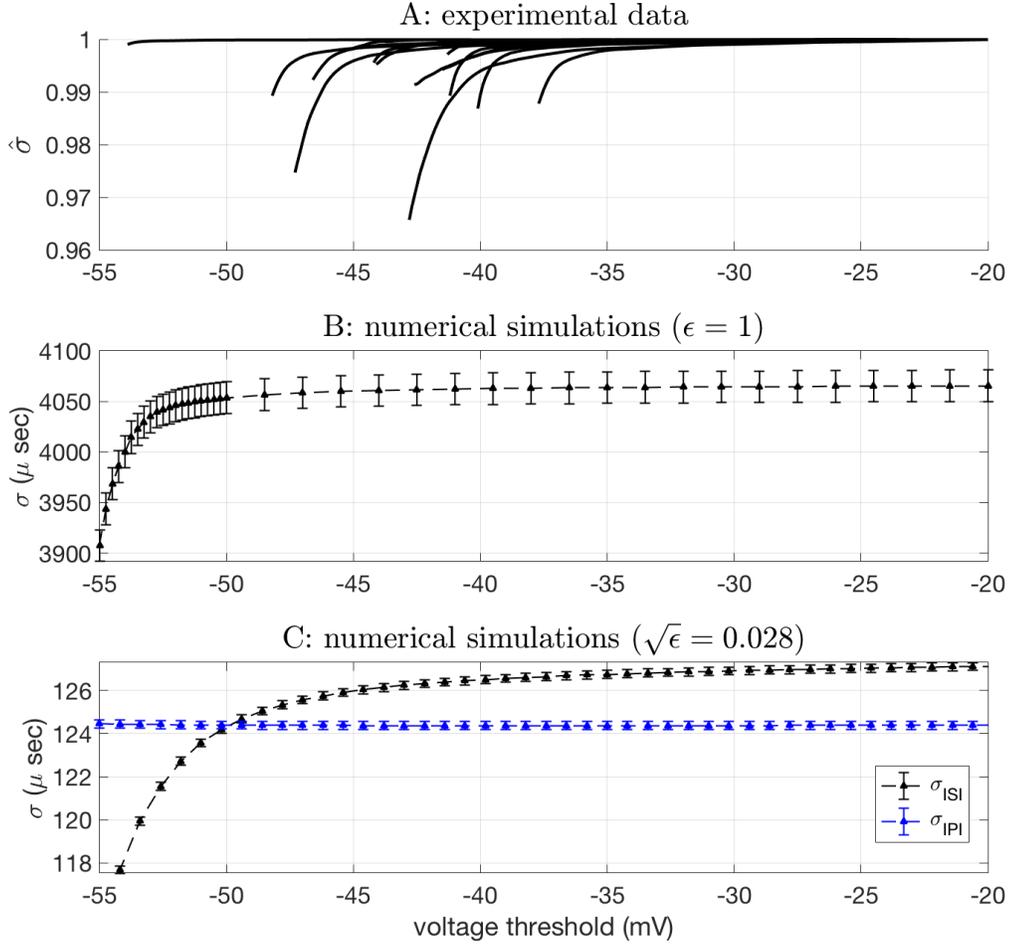


Figure 11.2: Standard deviation of the interspike intervals ($\sqrt{\sigma_{\text{ISI}}^2}$, cf. Def. 5) and standard deviation of the iso-phase intervals ($\sqrt{\sigma_{\text{IPI}}^2}$, cf. Def. 6) as a function of voltage threshold. **A:** Rescaled ISI standard deviation ($\hat{\sigma}$) obtained from experimental data recordings from 14 wild type Purkinje cells (data recorded in Dr. Friel’s laboratory). For each cell, $\sqrt{\sigma_{\text{ISI}}^2}$ is calculated using voltage threshold ranging from -55 mV to -20 mV, and scaled by dividing the stand deviation at voltage=-20 mV. **B, C:** standard deviation of ISI when $\epsilon = 1$ and $\sqrt{\epsilon} = 0.028$ in equation (9.28), respectively. For each voltage threshold, 500 different traces are generated with each trace containing roughly 1000 interspike intervals. Error bars indicate the 95% confidence interval of $\sqrt{\sigma_{\text{ISI}}^2}$ at each threshold. Note the vertical axis is in μ sec. In **C**, each value of σ_{IPI}^2 is calculated for the mean–return-time isochron intersecting the deterministic limit cycle at the voltage specified.

step of 0.008 ms, consisting of at least 600 ISIs, which was collected as one realization for the corresponding $\sigma_{\text{ISI}}(\epsilon, V_{\text{th}})$. The mean and standard deviation of the $\sigma_{\text{ISI}}(\epsilon, V_{\text{th}})$ is calculated for 1,000 realizations of the aforementioned step for each pair of ϵ and V_{th} . The error bars in Fig. 11.2 B and C indicate 95% confidence intervals of the standard deviation. As illustrated in Fig. 11.2 B and C, the standard deviation gradually increases as the trigger threshold increases during the AHP, and this trend is observed for both small and large noises. When $\epsilon = 1$, the noisy system in eqn. (9.28) is not close to the deterministic limit cycle, and there is not a good approximation for the phase response curve. When $\sqrt{\epsilon} = 0.028$, the system eqn. (9.28) can be considered in the small-noise region and thus we can find a corresponding phase on the limit cycle as the asymptotic phase. As shown in Fig. 11.2 C, unlike the variance of ISI, the variance of IPI is invariant with the choice of the phase threshold (ϕ).

11.2 Numerical Performance of the Decomposition Theorem

In this section, we will apply estimation methods based on Theorem 4 and Corollary 7 to the decomposition of variance of interspike intervals (ISIs, σ_{ISI}^2) and variance of inter-phase intervals (IPIs, σ_{IPI}^2), and numerically test their performance.

Fig. 11.3 presents a detailed comparison of the predicted and measured values of σ_{ISI}^2 and σ_{IPI}^2 , when the simulations only include noise from the K^+ channels. The channel noise generated by the Na^+ edges is suppressed by applying the stochastic shielding (SS) method to eqn. (9.28). For each plot in Fig. 11.3, 1000 repeated trials are collected and each trial simulates a total time of 15,000 milliseconds which generates more than 1000 ISIs or IPIs. Given our previous observation that σ_{ISI}^2 depends on the choice of voltage threshold, we

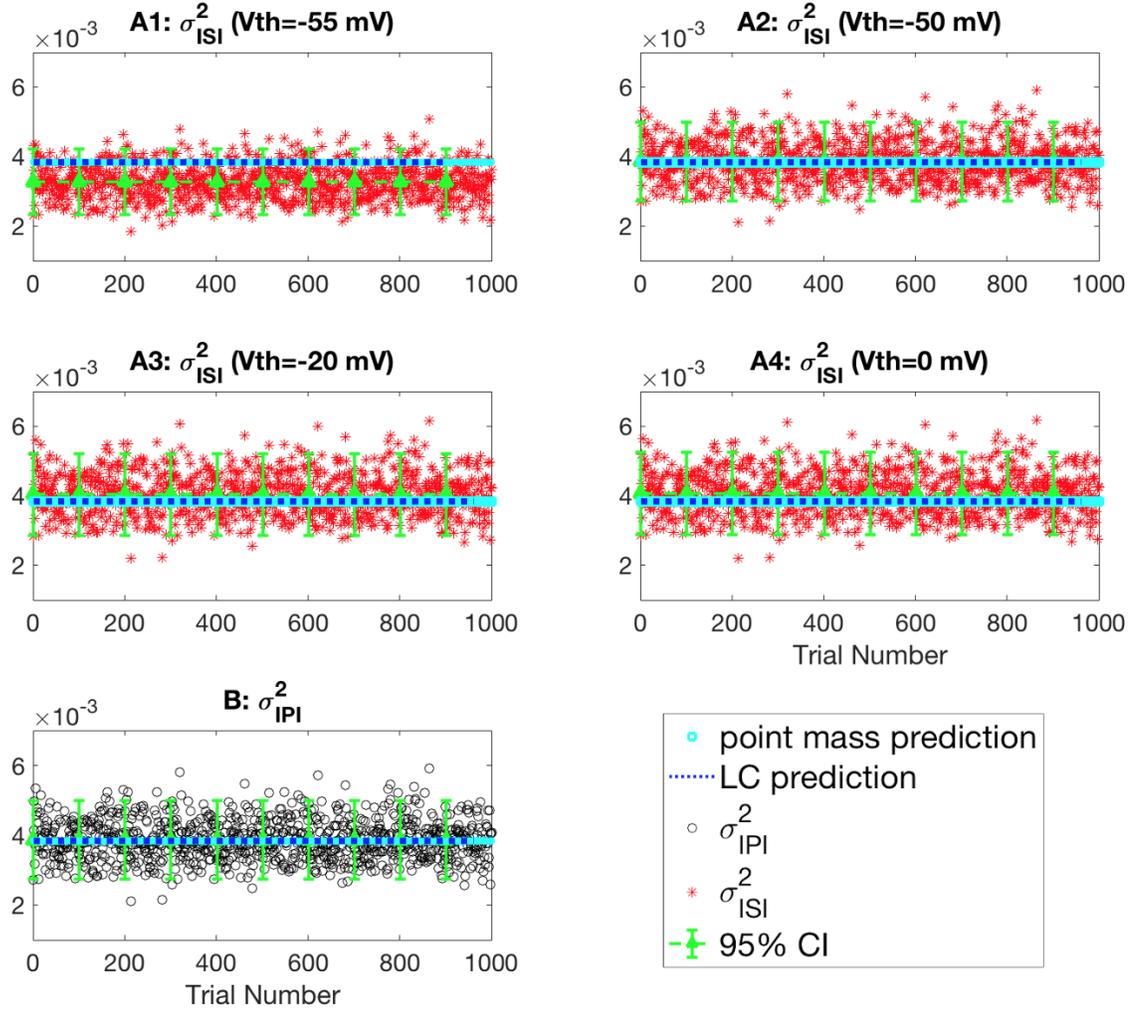


Figure 11.3: Variance of ISIs (σ_{ISI}^2) and IPIs (σ_{IPI}^2) with only K^+ edges included using the stochastic shielding method. Cyan dots: point mass prediction (cf. eqn. (11.2)). Solid blue line: limit cycle prediction. 1000 repeated simulations are plotted and for each of the simulation, more than 1000 ISIs (or IPIs) are recorded. Each sample point in the plot represents the variance of ISIs and IPIs for one realization. **A1-4**: Voltage threshold $V_{\text{th}} = -55, -50, -20, 0$ mV, with noise level $\sqrt{\epsilon} = 0.028$ (effective number of K^+ channels $N_{\text{tot}} \approx 2.30 \times 10^6$). **B**: Iso-phase section is the mean-return-time isochron intersecting the deterministic limit cycle at $V_{\text{th}} = -50$ mV.

selected four different voltage thresholds for comparison.

In Fig. 11.3, red dots in panels A1-A4 mark the ISI variance measured directly from simulated voltage traces, using the indicated V_{th} as the trigger voltage. Green stem-and-line marks show the mean and 95% confidence intervals of the direct ISI variance measure, calculated from all 1000 samples. The blue dotted line shows the ISI variance predicted from the limit cycle based estimate of the IPI variance (eq. (10.58)), and cyan squares show individual estimates using the point-mass prediction (eq. (11.3)). Note each point mass is an independent random variable; these estimates cluster tightly around the limit cycle based estimate. Panel B shows the variance of the inter-phase intervals calculated directly from the same 1000 trajectories (as described below), marked in black circles. Green stem-and-line marks show the mean and 95% C.I. for the IPI variance. The blue dotted line and cyan squares represent the same LC-based and point mass based IPI variance estimates as in A1-A4.

As shown in Fig. 11.3 (A1, A2, A3, A4 and B) the point mass prediction and the LC prediction of the IPI variance give almost the same result. Specifically, the LC prediction $\approx 3.84 \times 10^{-3}$ and the mean of the point mass predictions $\approx 3.83 \times 10^{-3}$ with a variance of $\approx 6.3 \times 10^{-11}$. Therefore, the LC prediction based on Corollary 7 gives a good approximation to the point mass prediction based directly on Theorem 4. For a given edge (or a group of edges) the LC prediction depends linearly on the scaling factor, ϵ , and can be easily calculated for various noise levels. Throughout the rest of this section, we will use the LC prediction as our predicted contribution from the decomposition theorem.

The asymptotic phase is calculated using equation (11.2) for each point on the stochastic trajectory. For a given voltage threshold, V_{th} , the corresponding iso-phase section is the mean-return-time isochron intersecting the deterministic limit cycle at V_{th} . As previously observed, the variance of the IPIs is invariant with respect to the choice of the reference iso-

phase section. As shown in Fig. 11.3 B, the prediction of variance of IPIs ($\approx 3.83 \times 10^{-3} \text{ ms}^2$) has a good match with the mean value of numerical simulations ($\approx 3.85 \times 10^{-3} \text{ ms}^2$). The 95% confidence interval of the IPIs are also plotted in Fig. 11.3 B, which further indicates the reliability of the prediction.

As shown in Fig. 11.3 (A1, A2, A3 and A4), with $V_{\text{th}} \in [-55, 0]$ mV, the numerical realizations of σ_{ISI}^2 are close to the predictions from the main theorem. However, the accuracy depends on the choice of the voltage threshold. As noted above, when the trigger voltage V_{th} is set below -50mV (for example, -55mV in Fig. 11.3,A1), the measured variance of ISIs falls below the value predicted from the IPI variance. When $V_{\text{th}} \approx -50\text{mV}$, the empirically observed value of σ_{ISI}^2 gives the best match to the IPI variance (cf. Fig. 11.2,C). When the trigger voltage V_{th} exceeds -50mV (for example, -20mV in Fig. 11.3,A3, and 0mV in Fig. 11.3,A4), the empirically observed variance of the ISIs is consistently higher than the IPI variance. Nevertheless, although the empirically observed numerical values of σ_{ISI}^2 ($\approx 4.00 \times 10^{-3} \text{ ms}^2$) overestimate the IPI-derived value when $V_{\text{th}} > -50\text{mV}$, they remain close to the IPI value. Fig. 11.3 panels A1-4 show that even though the IPI-based prediction of the ISI variance works best when the trigger voltage is set to $V_{\text{th}} \approx -50\text{mV}$, the IPI-based variance falls within the 95% confidence interval of σ_{ISI}^2 regardless of the value of V_{th} chosen. Therefore, we can conclude that Theorem 4 and Corollary 7 give a good approximation to the value of σ_{ISI}^2 , at least at noise level $\sqrt{\epsilon} = 0.028$.

Practically, the voltage-based interspike interval variance, σ_{ISI}^2 , is a more widely used quantity [57, 74, 86, 112, 118] because it can be calculated directly from electrophysiological recordings. The inter-phase interval variance, σ_{IPI}^2 , however, can not be directly measured or calculated. Even given the stochastic model with its realizations, calculating the asymptotic phase and finding the IPIs are numerically expensive. Despite its lack of consistency, as shown in Fig. 11.3 (A3 and A4), the σ_{ISI}^2 can approximately be decom-

posed using Theorem 4 and Corollary 7, which offer predicted values that fall in the 95% confidence interval of σ_{ISI}^2 .

Fig. 11.4 summarizes the overall fit of the decomposition of variance of ISIs to the prediction from Theorem 4 and Corollary 7. We applied the stochastic shielding method by including each directed edge separately in the transition graph (cf. Fig. 3.1). In Fig. 11.4 (B and D), the variance of the ISIs is compared with the value obtained with the limit cycle based prediction from eqn. (10.58).

Fig. 11.4 (A and C) shows the log-log plot for the ISI variance (σ_{ISI}^2) of each individual edge as a function of the noise level, ϵ , in the range of $[e^{-10}, e^5]$, measured via direct numerical simulation using $V_{\text{th}} = -20$ mV. The color for each edge ranges from red to blue according an ascending order of edge numbers (1-8 for K^+ and 1-20 for Na^+). The total effective number of Na^+ channels is $M_{\text{tot}} = M_{\text{ref}}/\epsilon$ and of K^+ channels is $N_{\text{tot}} = N_{\text{ref}}/\epsilon$, where the reference channel numbers are $M_{\text{ref}} = 6000$ and $N_{\text{ref}} = 1800$ (described in §9.4). That is, we consider ranges of channel numbers $40 \lesssim M_{\text{tot}} \lesssim 1.3 \times 10^8$ for Na^+ and $12 \lesssim N_{\text{tot}} \lesssim 4.0 \times 10^7$ for K^+ . Thus, we cover the entire range of empirically observed single-cell channel populations (cf. Tab. 3.1).

As shown in Fig. 11.4 (A and C), the linear relation between σ_{ISI}^2 and ϵ predicted from Theorem 4 is numerically observed for all 28 directed edges in the Na^+ and K^+ transition graphs (cf. 3.1) for small noise. The same rank order of edge importance discussed in §6.5 is also observed here in the small noise region. Moreover, the smaller the edge importance measure for an individual edge, the larger the value of ϵ before observing a breakdown of linearity.

Fig. 11.4 (B and D) presents the log-log plot for the ISI variance (σ_{ISI}^2 , black solid line) when including noise only from the K^+ edges and Na^+ edges, respectively. As in panels A and C, the noise level, ϵ is in the range of $[e^{-10}, e^5]$. The LC prediction for eqn. (10.58)

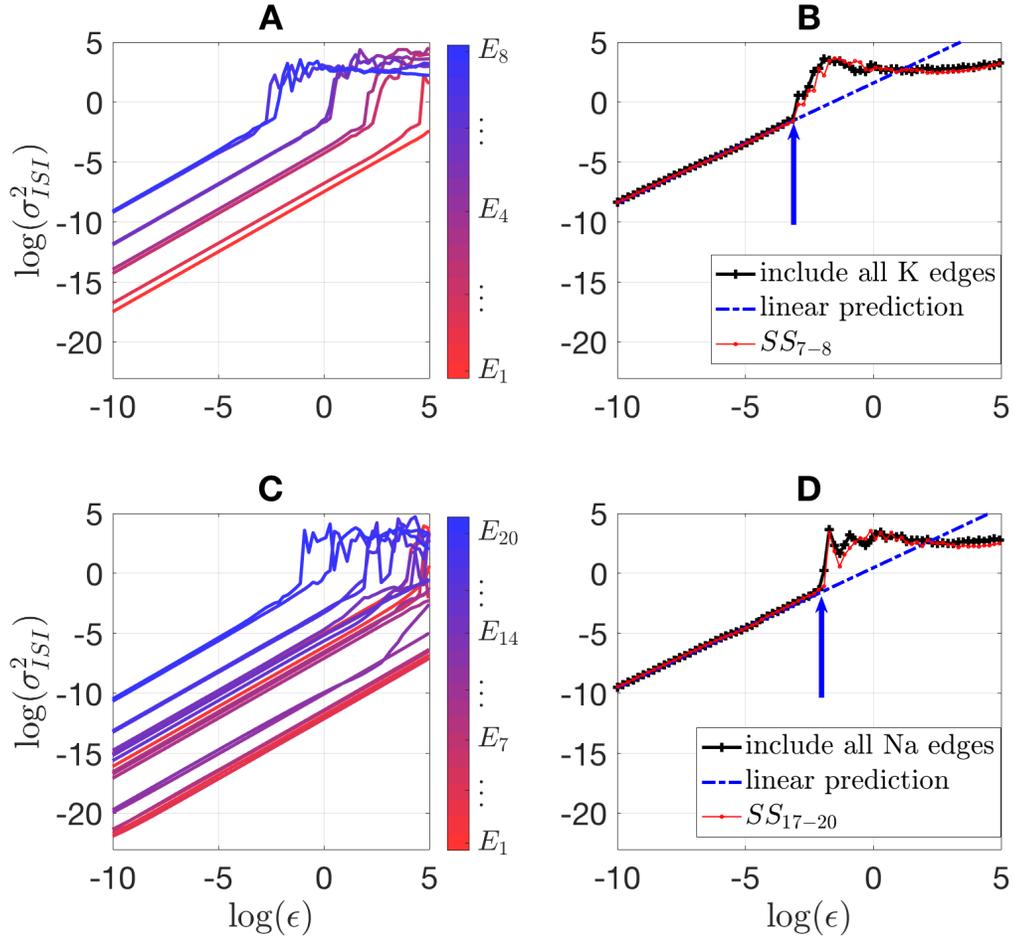


Figure 11.4: Numerical performance of the decomposition for the ISI variance for Na^+ and K^+ kinetics. **A&C:** σ_{ISI}^2 for individual K^+ (**A**) and Na^+ (**C**) edges. E_k marks the k th edge, $k \in \{1, \dots, 8\}$ for K^+ and $k \in \{1, \dots, 20\}$ for Na^+ . **B&D:** Linearity of superposition for K^+ (**B**) and Na^+ (**D**) channels. See text for details. σ_{ISI}^2 with noise from all K^+ (Na^+) edges included (black line), with only the most significant edges using stochastic shielding (red line) included, and the linear prediction from Theorem 4 (blue dashed line).

from Theorem 4 when including noise from only the K^+ (or Na^+) channels is plotted in dashed blue. For example, the linear noise prediction for the potassium channels alone is

$$\sigma_{\text{ISI}}^2 \approx \sum_{k=1}^{\mathcal{E}_K} \sigma_{\text{ISI},k}^2 \quad (11.4)$$

where $\mathcal{E}_K = 8$ (similarly, $\mathcal{E}_{Na} = 20$), and $\sigma_{\text{ISI},k}^2$ is the LC prediction for the k^{th} edge. As shown in Fig. 11.4 panel B, the linear prediction matches well with the numerically calculated σ_{ISI}^2 up to $\ln(\epsilon) \approx -3.0$ (indicated by the blue arrow) which corresponds to approximately 36,000 K^+ channels. For Na^+ , the theorem gives a good prediction of the numerical σ_{ISI}^2 up to $\ln(\epsilon) \approx -1.9$ (indicated by the blue arrow) which corresponds to approximately 40,000 Na^+ channels. These channel population sizes are consistent with typical single-cell ion channel populations, such as the population of Na^+ channels in the node of Ranvier, or the Na^+ and K^+ channels in models of the soma of a cerebellar Purkinje cell (cf. Tab. 3.1).

Finally, we apply stochastic shielding (SS) to both the K^+ and Na^+ channels by only including noise from the edges making the largest contributions in Fig. 11.4 panels A and C. For the K^+ channel, we include edges 7 and 8, and for Na^+ , we include edges 17, 18, 19 and 20. As shown in Fig. 11.4 panels B and D, the SS method (solid red line) gives a good match to the overall σ_{ISI}^2 for all noise intensities $\epsilon \in [e^{-10}, e^5]$, with numbers of K^+ channels ≥ 12 and Na^+ channels ≥ 40 .

Fig. 11.5 shows the overall performance of the prediction of σ_{ISI}^2 based on Theorem 4, when noise from all 28 directed edges are included (black line). The theorem is stated as an asymptotic result in the limit of weak noise. The predicted ISI variance using the theorem (dashed blue curve) matches the ISI variance obtained from the full numerical simulation for modest noise levels, up to $\ln(\epsilon) \leq -3.9$, corresponding to $\gtrsim 90,000$ K^+ channels and

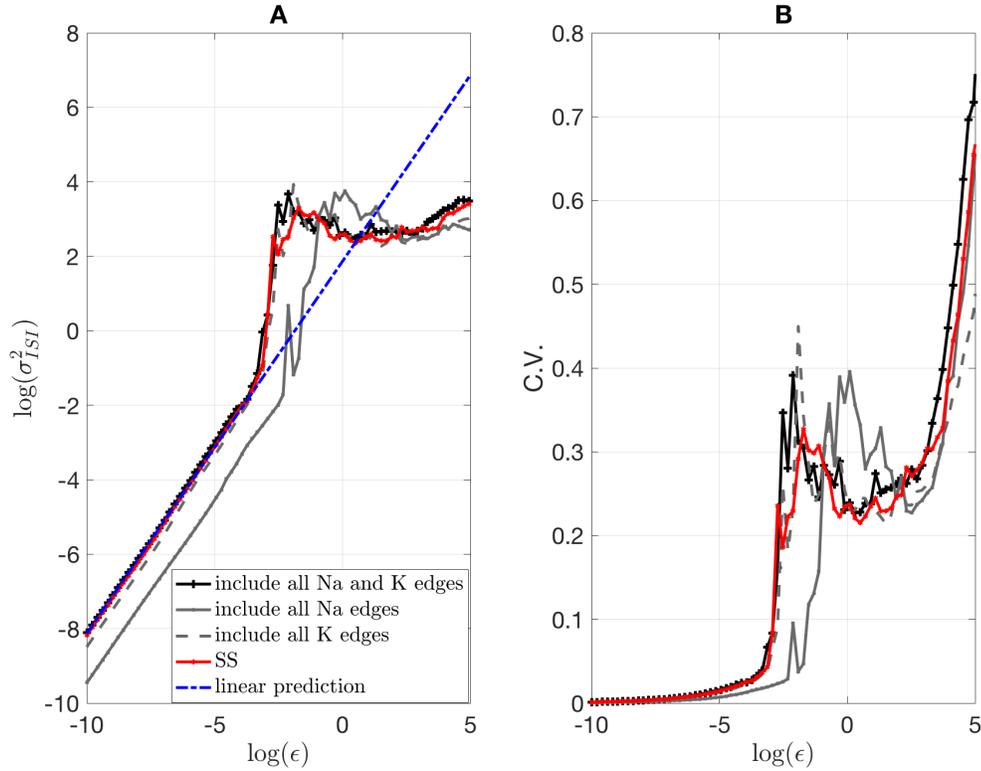


Figure 11.5: Numerical performance of the decomposition for the ISI variance of the full system. **A:** Log-log plot of σ_{ISI}^2 for $\epsilon \in [e^{-10}, e^5]$. ISI variance contribution σ_{ISI}^2 with noise from all 28 edges included (black line), only from 8 K^+ edges (dashed grey), only from 20 Na^+ edges (solid grey), and SS using noise from six edges (red line, see text for details). The linear prediction from Theorem 4 for the whole system is plotted for comparison (dashed blue line). **B:** Coefficient of variation (C.V.), or mean ISI divided by $\sqrt{\sigma_{ISI}^2}$, vs. $\log(\epsilon)$. Same color scheme as **A**. Compare Fig. 4.1, which shows data from two cerebellar Purkinje cells, a wild-type cell with C.V. ≈ 0.039 and a cell from a *leaner* mouse with C.V. ≈ 0.30 .

$\gtrsim 300,000$ Na^+ channels. These population sizes are at the high end of the range of typical numbers of channels neurons (cf. Tab. 3.1).

For smaller ion channel populations (larger noise levels), the linear approximation breaks down, but the stochastic shielding approximation remains in good agreement with full numerical simulations. Fig. 11.5 shows σ_{ISI}^2 from simulations using the SS method including only noise from the six most important edges (edges 7-8 in K^+ and 17-20 in Na^+), plotted in solid red. For $\ln(\epsilon) \gtrsim -3.5$, both the full simulation and the SS simulation show a rapid increase in σ_{ISI}^2 with increasing noise level. This dramatic increase in timing variability results when increasing noise causes the neuron to “miss” spikes, that is, to generate a mixture of regular spiking and small subthreshold oscillations [101]. Including noise from all 20 Na^+ channel edges (gray line) or all eight K^+ channel edges (gray dashed line) shows a similar jump, albeit delayed to higher values of ϵ for the Na^+ channel. Note also the Na^+ channel alone has a quantitatively smaller contribution to ISI variability for the stochastic HH model than the K^+ channel for all noise levels in the linear region.

For larger noise levels ($\ln(\epsilon) \gtrsim -2$), all simulations become sufficiently noisy that $\ln(\sigma_{\text{ISI}}^2)$ collapse to a similar level, approximately 3. As the interspike interval is a non-negative random quantity with a constrained mean (bounded by the reciprocal of the firing rate), once the spike train has maximal variability, further increasing the strength of the channel noise does not drive the ISI variance appreciably higher. However, although the ISI variance appears approximately to saturate with increasing noise, the coefficient of variation (C.V., $\sqrt{\sigma_{\text{ISI}}^2}/I$) continues to increase (Fig. 11.5B), because the mean ISI (I) decreases with increasing noise (the firing rate increases with increasing noise, data not shown).

Part V

Conclusions and Discussion

Chapter 12

Conclusions

12.1 Summary

The exact method for Markov chain (MC) simulation for an electrotonically compact (single compartment) conductance-based stochastic model under current clamp is a hybrid discrete (channel state) / continuous (voltage) model of the sort used by [4, 18, 87]. While MC methods are computationally expensive, simulations based on Gaussian/Langevin approximations can capture the effects of stochastic ion channel fluctuations with reasonable accuracy and excellent computational efficiency. Since Goldwyn and Shea Brown's work focusing the attention of the computational neuroscience community on Fox and Lu's Langevin algorithm for the Hodgkin-Huxley system [39, 49], several variants of this approach have appeared in the literature.

In this thesis, as in [95], we advocate for a class of models combining the best features of conductance-based Langevin models with the recently developed stochastic shielding approximation [102, 104, 105]. We propose a Langevin model with a 14-dimensional

state space, representing the voltage, five states of the K^+ channel, and eight states of the Na^+ channel; and a 28-dimensional representation of the driving noise: one independent Gaussian noise term for each directed edge in the channel-state transition graph. We showed in §5.3 that the corresponding mean-field 14D ordinary differential equation model is consistent with the classical HH equations in the sense that the latter correspond to an invariant submanifold of the higher-dimensional model, to which all trajectories converge exponentially quickly. Fig. 5.1 illustrated the relation between the deterministic 4D and 14D Hodgkin Huxley systems. Building on this framework, we introduced the $14 \times 28D$ model, with independent noise sources corresponding to each ion channel transition (§6). We proved in §7 that, given identical boundary conditions, our $14 \times 28D$ model is pathwise equivalent both to Fox and Lu’s original Langevin model, and to a 14-state model with 14 independent noise sources due to [90].

The original 4D HH model, the 14D deterministic HH model, and the family of equivalent 14D Langevin models we consider here, form a nested family, each contained within the next. Specifically, (i) the 14D ODE model is the “mean-field” version of the 14D Langevin model, and (ii) the 4D ODE model forms an attracting invariant submanifold within the 14D ODE model, as we establish in our Lemma 2. So in a very specific sense, the original HH equations “live inside” the 14D Langevin equations. Thus these three models enjoy a special relationship. In contrast, the widely used 4D Langevin equations, originally studied in [40], do not bear an especially close relationship to the other three.

In addition to rigorous mathematical analysis we also performed numerical comparisons (§8) showing that, as expected, the pathwise equivalent models produced statistically indistinguishable interspike interval (ISI) distributions. Moreover, the ISI distributions for our model (and its equivalents) were closer to the ISI distribution of the “gold standard” MC model under two different metric space measures. Our method (along with Orio and

Soudry's) proved computationally more efficient than Fox and Lu's original method and Dangerfield's model [21]. In addition, our method lends itself naturally to model reduction (via the stochastic shielding approximation) to a significantly faster $14 \times 6D$ simulation that preserves a surprisingly high level of accuracy.

We prove in §10 that the numerically calculated edge importance can be quantified from the molecular-level fluctuations of the stochastic Hodgkin-Huxley (HH) kinetics. Specifically, we combine the stochastic shielding approximation with the re-scaled Langevin models (eqn. (9.28)) of the HH model to derive analytic results for decomposing the variance of the cycle time (the iso-phase intervals) for mean–return-time isochrons of the stochastic HH models. We prove in theory, and confirm via numerical simulations, that in the limit of small noise, the variance of the iso-phase intervals decomposes linearly into a sum of contributions from each edge. We show numerically that the same decomposition affords an efficient and accurate estimation procedure for the interspike intervals, which are experimentally observable. Importantly, our results apply to current clamp rather than to voltage clamp conditions. Under current clamp, a stochastic conductance-based model is an example of a piecewise-deterministic Markov process (PDMP). We show in §11.2 that our theory is exact in the limit of small channel noise (equivalently, large ion channel population size); through numerical simulations we demonstrate its applicability even in a range of small to medium noise levels, consistent with experimentally inferred single-cell ion channel population sizes. In addition, we present the numerical performance of the SS method under different scenarios and argue that the stochastic-shielding approximation together with the $14 \times 28D$ Langevin representation give an excellent choice of simulation method for ion channel populations spanning the entire physiologically observed range.

Chapter 13

Discussions

13.1 Discrete Gillespie Markov Chain Algorithms

The discrete-state Markov chain algorithm due to Gillespie is often taken to be the gold standard simulation for a single-compartment stochastic conductance-based model. Most former literature on Langevin HH models, such as [49, 60, 73, 90], when establishing a reference MC model, consider a version of the discrete Gillespie algorithm that assumes a piecewise-constant propensity approximation, i.e. that does not take into account that the voltage changes between transitions, which changes the transition rates. This approximation can lead to biophysically unrealistic voltage traces for very small system sizes (cf. Fig. 2 of [67], top trace with $N = 1$ ion channel) although the differences appear to be mitigated for $N \gtrsim 40$ channels [4]. In the present thesis, our MC simulations are based on 6000 Na^+ and 1800 K^+ channels (as in [49]), and we too use the ISI distribution generated by a piecewise-constant propensity MC algorithm as our reference distribution. Given the range of physiologically observed single-cell ion channel population sizes (Tab. 3.1), it ap-

pears the piecewise-constant-propensity approximation is sufficient for naturally occurring cells.

However, as shown in Tab. 8.1 and Fig. 8.2, the computation time for the MC simulations is one order of magnitude larger than efficient methods such as [21, 90] and the $14 \times 28\text{D}$ model. The computational cost of the MC model increases dramatically as the number of ion channels grows, therefore, even the approximate MC algorithm is inapplicable for a large number of channels.

13.2 Langevin Models

It is worth pointing out that the accuracy of Fox and Lu’s original Langevin equations has not been fully appreciated. In fact, Fox and Lu’s model [39] gives an approximation to the MC model that is just as accurate as [90] both in the gating variable statistics [49] and also in the ISI distribution sense (see §8) – because as we have established, these models are pathwise equivalent! However, the original implementation requires taking a matrix square root in every timestep in the numerical simulation, which significantly reduces its computational efficiency.

Models based on modifications of [39]’s work can be divided into three classes: the subunit model [40]; effective noise models [56, 73], and channel-based Langevin models such as [21, 49, 60, 90, 93].

Subunit model The first modification of Fox and Lu’s model is the subunit model [40], which keeps the original form of the HH model, and adds noise directly to the gating variables (m , h , and n) [49, 40]. The subunit approximation model was widely used because of its computational speed. However, as [12] and others pointed out, the inaccuracy

of this model remains significant even for large number of channels. Moreover, [49] and [61] found that the subunit model fails to capture the statistics of the HH Na⁺ and K⁺ gates. In this thesis, we also observed that the subunit model is more efficient than channel-based Langevin models, but tends to delay spike generation. As shown in Fig. 8.1, the subunit model generates significantly longer ISIs than the MC model.

Effective noise models Another modification to Fox and Lu’s algorithm is to add colored noise to the channel open fractions. We did not include colored noise models such as [73, 56] in our model comparison. However, [61] found that both these effective noise models generate shorter ISIs than the MC model with the same parameters. Though the comparison we provided in §8 only include the Fox and Lu 94, Fox97, Goldwyn, Dangerfield, Orio, SS and the $14 \times 28D$ model, combining the results from [49] and [61], the $14 \times 28D$ model could be compared to a variety of models including [21, 39, 40, 49, 56, 60, 73, 90].

Channel-based Langevin models The main focus of Part III (§7-§8) of this thesis is the modification based on the original Fox and Lu’s matrix decomposition method, namely, the channel-based (or conductance-based) Langevin models. We proved in §7 that under the same boundary conditions, Fox and Lu’s original model, Orio’s model and our $14 \times 28D$ model are pathwise equivalent, which was also verified from our numerical simulations in §7 and §8. In §7, we discussed channel-based Langevin models including [21, 39, 41, 49, 90]. We excluded Fox’s more recent implementation [41] in §8 for two reasons. First, the algorithm is pathwise equivalent to others considered there. And moreover, the method is vulnerable to numerical instability when performing the Cholesky decomposition. Specifically, some of the elements in the S matrix from the Cholesky decomposition in [41] involve square roots of differences of several quantities, with no guarantee that the

differences will result in nonnegative terms – even with strictly positive values of the gating variables. Nevertheless, this model would be in the equivalence class and in any case would not be more efficient than Orio’s model, because of the noise dimension and complicated operations (e.g. taking multiple square roots) in each time step.

13.3 Model Comparisons

If two random variables have similar distributions, then they will have similar moments, but not necessarily vice-versa. Therefore, comparison of the full interspike-interval distributions produced by different simulation algorithms gives a more rigorous test than comparison of first and second moments of the ISI distribution. Most previous evaluations of competing Langevin approximations were based on the accuracy of low-order moments, for example the mean and variance of channel state occupancy under voltage clamp, or the mean and variance of the interspike interval distribution under current clamp [21, 48, 49, 60, 61, 73, 90, 102]. In this thesis, we compare the accuracy of the different algorithms using the full ISI distribution, but using the L_1 norm of the difference (Wasserstein metric) and the L_∞ norm (Kolmogorov-Smirnov test). The paper [52] previously compared the ISI distributions generated by the Markov chain (Gillespie algorithm) to the distribution generated by different types of Langevin approximations (LA), including the original Langevin models [39, 49], the channel-based LA with colored noise [56, 73], and LA with a 14×14 variant of the diffusion coefficient matrix S [90]. They concluded that Orio and Soudry’s method provided the best match to the Markov chain model, specifically “Fox-Goldwyn, and Orio-Kurtz¹ methods both generate ISI histograms very close to those of Micro²” [52]. We note that the comparison reported in this paper simply superimposed plots of the ISI

¹We refer to this model as to as “Orio”

²This is the model we refer to as the Markov-chain model.

distributions, allowing a qualitative comparison, while our metric-space analysis is fully quantitative. In any case, their conclusions are consistent with our findings; we showed in §7 that the Fox-Goldwyn and the Orio-Kurtz model are pathwise equivalent (when implemented with the same boundary conditions), which accounts for the similarity in the ISI histograms they generate. In fact, because of pathwise equivalence, we can conclude that the true distributions for these models are identical, and any differences observed just reflect finite sampling.

13.4 Stochastic Shielding Method

The stochastic shielding (SS) approximation [102] provides an efficient and accurate method for approximating a Markov process using only a subset of transitions to generate fluctuations, namely those directly involving observable states. For conductance-based models, rather than aggregating ion channel states, SS effects dimension reduction by selectively eliminating those independent noise sources (channel state transitions) that have the least impact on current fluctuations. Recent work in [93] compared previous methods such as [21, 46, 60, 90, 102] in accuracy, applicability and simplicity as well as computational efficiency. They concluded that for mesoscopic numbers of channels, stochastic shielding methods combined with diffusion approximation methods can be an optimal choice. That is precisely the combination of methods that we advocate in this thesis.

Like [90], the stochastic shielding method proposed by [93] assumed detailed balance of transitions between adjacent states and used edges that are directly connected to the open gates of HH Na^+ and K^+ . We calculated the edge importance in §6.5 and found that the four (out of twenty) most important directed edges for the Na^+ gates are *not* the four edges directly connected to the conducting state, as assumed in previous application of the SS

method [102], but rather the two pairs of edges “upstream” of the open state, which are the transitions typically leading to initiation of a voltage spike.

13.5 Which Model to Use?

Among all the modifications of Fox and Lu’s method considered here, Orio and Soudry’s approach, and our $14 \times 28D$ model, provide the best approximation to the “gold standard” MC model, with the greatest computational efficiency. Several earlier models were studied in the review paper by [49], where they rediscovered that the original Langevin model proposed by Fox and Lu is the best approximation to the MC model among those considered. Later work ([61]) further surveyed a wide range of Langevin approximations for the HH system including [39, 40, 49, 56, 60, 73, 90] and explored models with different boundary conditions. The survey [61] concluded that the bounded and truncated-restored Langevin model [60] and the unbounded [90]’s model provide the best approximation to the MC model.

As shown in §7 and §8, the $14 \times 28D$ Langevin model naturally derived from the channel structure is pathwise equivalent to the Fox and Lu ‘94, Fox ‘18, and the Orio-Soudry models under the same boundary conditions. The $14 \times 28D$ model (with open boundary conditions) is more accurate than models with the reflecting boundary condition method of [21], and also better than the approximation method proposed by [49], when the entire ISI distribution is taken into account. We note that [61, 111] treated Goldwyn’s method [49] as the original Fox and Lu model in their comparison, however, the simulation in [49] uses the $4D$ multinomial submanifold to update the gating variables. Our analysis and numerical simulations suggest that the original Fox and Lu model is indeed as accurate as the Orio-Soudry model, while the computational cost still remains a major concern.

Though the $14 \times 28\text{D}$ model has similar efficiency and accuracy with [90], it has several advantages. First, the rectangular S matrix (in eqn. (3.2)-(3.3)) in Orio's model merges the noise contributions of reciprocal pairs of edges. However, this dimension reduction assumes, in effect, that detailed balance holds along reciprocal edges, which our results show is not the case, under current clamp (Fig. 6.2). Moreover, the $14 \times 28\text{D}$ model arises naturally from the individual transitions of the exact evolution equations (eqn.(6.14)-(6.15)) for the underlying Markov chain model, which makes it conceptually easier to understand. In addition, our method for defining the $14 \times 28\text{D}$ Langevin model and finding the best SS model extends to channel-based models with arbitrary channel gating schemes beyond the standard HH model. Given any channel state transition graph, the Langevin equations may be read off from the transitions, and the edge importance under current clamp can be evaluated by applying the stochastic shielding method to investigate the contributions of noise from each individual directed edge. Finally, in exchange for a small reduction in accuracy, the stochastic shielding method affords a significant gain in efficiency. The $14 \times 28\text{D}$ model thus offers a natural way to quantify the contributions of the microscopic transitions to the macroscopic voltage fluctuations in the membrane through the use of stochastic shielding. For general ion channel models, extending a biophysically-based Langevin model analogous to our $14 \times 28\text{D}$ HH model, together with the stochastic shielding method, may provide the best available tool for investigating how unobservable microscopic behaviors (such as ion channel fluctuations) affect the macroscopic variability in many biological systems.

13.6 Variance of Interspike Intervals

There are several different methods for detecting spikes and quantifying interspike intervals (ISIs). In one widely used approach [44, 57, 74, 84, 86, 112, 118], we can define the threshold as the time of upcrossing a fixed voltage, which is also called a Schmitt trigger (after O.H. Schmitt [106]). We primarily use this method in this thesis.

As an alternative, the time at which the rate of change of voltage, dV/dt , reaches its maximum value (within a given spike) has also been used as the condition for detecting spikes [7]. However, in contrast with the voltage-based Schmitt trigger, using the maximum of dV/dt to localize the spike does not give a well-defined Poincaré section. To see this, consider that for a system of the form (7.6) we would have to set

$$\begin{aligned} \frac{d^2V}{dt^2} &= \frac{d}{dt}f(V, \mathbf{N}) \\ &= f(V, \mathbf{N}) \frac{\partial f}{\partial V}(V, \mathbf{N}) - \frac{dM_8}{dt} g_{\text{Na}}(V - V_{\text{Na}}) - \frac{dN_5}{dt} g_{\text{K}}(V - V_{\text{K}}) \end{aligned} \quad (13.1)$$

equal to zero to find the corresponding section. The difficulty is evident: for the Langevin system the open fraction M_8 (resp. N_5) of sodium (resp. potassium) channels is a diffusion process, and is not differentiable, so “ dM_8/dt ” and “ dN_5/dt ” are not well defined. Moreover, even if we could interpret these expressions, the set of voltages V and gating variables \mathbf{N} for which (13.1) equals zero depends on the instantaneous value of the noise forcing, so the corresponding section would not be fixed within the phase space. For a discrete state stochastic channel model, the point of maximum rate of change of voltage could be determined post-hoc from a trajectory, but again depends on the random waiting times between events, and so is not a fixed set of points in phase space. For these reasons we do not further analyze ISIs based on this method of defining spikes, although we nevertheless

include numerical ISI variance based on this method, for comparison (see Fig. 13.1 below).

As a third possibility, used for example in [57], one sets the voltage nullcline ($dV/dt = 0$), at the top of the spike, as the Poincaré section for spike detection. That is, one uses a surface such as $\mathcal{S}_{\text{peak}} = \{(v, \mathbf{n}) \mid f(v, \mathbf{n}) = 0\} \cap \{v > -40\}$. This condition does correspond to a well-defined Poincaré section, albeit one with a different normal direction than the voltage-based sections.

In contrast to the ISI variance, which depends to some degree on the choice of spike-timing method used, the *mean* ISI is invariant. Both in numerical simulations and from experimental recordings, the mean interspike interval using any of the three methods above is very stable. But the apparent ISI variance changes, depending on the method chosen.

We observe in both real and simulated voltage traces that the ISI variance, σ_{ISI}^2 , depends not only on the method for identifying spikes, but also on the voltage used for the Schmitt trigger. To our knowledge this sensitivity of ISI variance to trigger voltage has not been previously reported. Generally speaking, from analyzing both simulation and recorded data from *in vitro* studies, the ISI variance is not a constant, but increases slightly as the voltage threshold defining a “spike” is increased (cf. §11.1). Thus the ISI variance is not an intrinsically precisely invariant quantity for model or real nerve cells.

Fig. 13.1 shows σ_{ISI}^2 obtained empirically from electrophysiological recordings of Purkinje cells *in vitro* (upper plot) and from simulations of the stochastic Hodgkin-Huxley system (lower plot) with a small noise amplitude ($\sqrt{\epsilon} = 0.028$) using the three methods for spike time extraction described above, for a single voltage trace comprising 785 interspike intervals. The ISI variance as a function of trigger voltage increases steadily from below 7.8 ms^2 to above 7.9 ms^2 as the trigger voltage increases from -50 mV to -20 mV . In contrast, the ISI variance obtained from the peak voltage ($dV/dt \approx 0$, obtained using linear interpolation of the first-order voltage difference) or the maximum slope condition

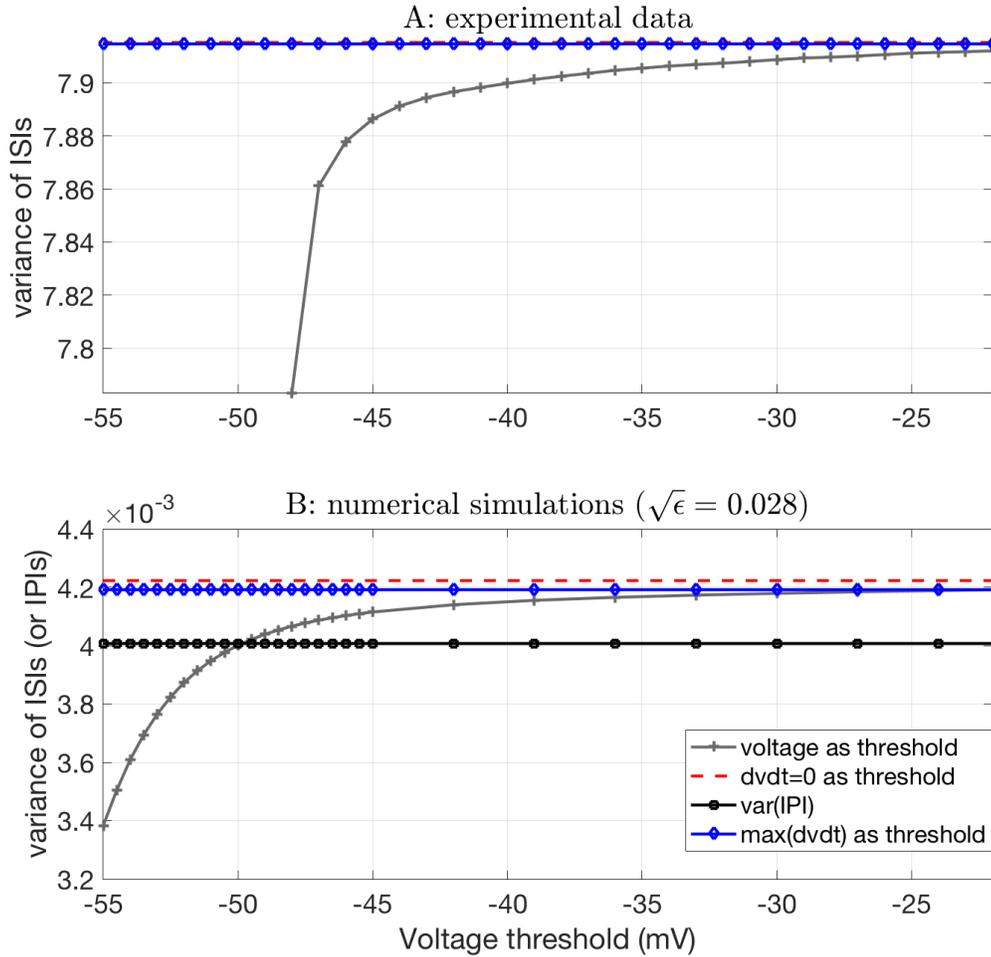


Figure 13.1: Variance of interspike intervals using different threshold conditions. **A:** σ_{ISI}^2 of spikes from a single trace of a wild type Purkinje cells comprising 785 ISIs. $\sigma_{\text{ISI}}^2 = 7.9151$ when setting $dv/dt = 0$ as the threshold (dashed red), and $\sigma_{\text{ISI}}^2 = 7.9146$ when maximum dv/dt is set to be the threshold condition (blue). Different voltage thresholds show increasing σ_{IS}^2 with voltage (gray). **B:** ISI variance from a Langevin HH (cf. eq. 9.28) simulation with small noise ($\sqrt{\epsilon} = 0.028$) comprising c. 1000 ISIs. Labels as in A. The variance of the inter-phase intervals is constant regardless of the particular isochron chosen (black).

($d^2V/dt^2 \approx 0$ and $dV/dt > 0$, obtained using linear interpolation of the second-order voltage difference) give nearly indistinguishable values (red and blue superimposed traces in Fig. 13.1A) that lie slightly above the largest value of σ_{ISI}^2 at the upper range of the trigger voltage.

A similar phenomenon occurs for Langevin simulations of the HH model with small noise (Fig. 13.1B). In this case, the ISI variance based on maximum slope falls slightly below the variance based on the spike peaks; both are similar to the variance obtained with a Schmitt trigger close to -20 mV. This similarity at higher trigger voltages probably occurs because the inflection point of each spike occurs at nearly the same voltage (at least, for small noise).

As shown in §10.2 and §11.2, the inter-phase interval (IPI, also referred as iso-phase interval), based on the crossing time of iso-phase sections, provides a uniform σ_{IPI}^2 for all choices of reference iso-phase sections (cf. Fig. 11.3). Fig. 13.1B shows the IPI variance (in black) for different mean–return-time isochronal sections, each passing through the limit cycle trajectory at the specified voltage.

For experimental voltage recordings, we cannot specify the interphase variables without a measurement or estimate of the entire state vector. Fortunately, the sensitivity of ISI variance to voltage threshold, while statistically significant, is relatively small (a few percent), as voltage is the practical measure available for marking spike times. Moreover, as shown §11.2, Theorem 4 and Corollary 7 can be is well suited to approximating the variance of ISIs (σ_{ISI}^2) despite its threshold-dependence.

For moderate to large noise Langevin model traces ($\epsilon \approx 1$), we also see a systematic shift in σ_{ISI}^2 with increasing Schmitt-trigger voltage. However, the size of the shift is an order of magnitude smaller than the variability of the variance across trials. Fig.13.2 plots σ_{ISI}^2 versus trigger voltage, as well as the ISI variance based on the peak voltage and the

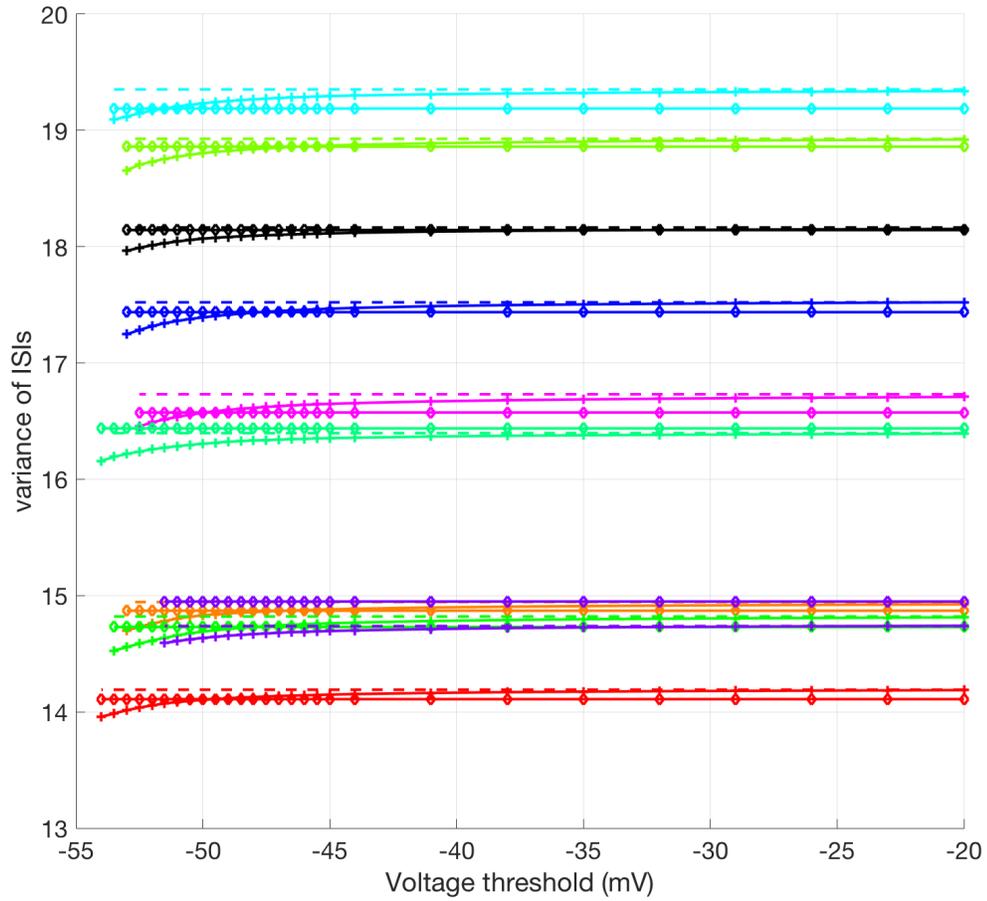


Figure 13.2: Model simulation of σ_{ISI}^2 for the Langvin model (eqn. 9.28) using different thresholds when $\epsilon = 1$. Ten repeated simulations are plotted, with each containing roughly 1000 ISIs. Solid plus: σ_{ISI}^2 using different voltages as threshold. Dashed: $dV/dt = 0$ as the spike condition. Diamonds: maximal dV/dt condition. Each color represents a different sample with independent noise. Variance is in units of ms^2 .

maximal slope conditions, for ten different samples of the Langevin HH model with $\epsilon = 1$, each comprising $\gtrsim 1000$ interspike intervals. In each case σ_{ISI}^2 is a smoothly increasing function of the trigger voltage, but the range of the increase in variance is approximately 0.25 ms^2 , while the sample variance of the ISI variance itself is approximately 3.5 ms^2 across the ten trials, an order of magnitude larger. For comparison, the sample variance of σ_{ISI}^2 across c. 4000 trials, cf. Fig. 11.3, is approximately $3 \times 10^{-7} \text{ ms}^2$. The source of the variance for the larger noise value may involve the introduction of extra or missing spikes from the regular spiking pattern, cf. Fig. 11.5. Thus, although σ_{ISI}^2 based on the standard Schmitt trigger approach is sensitive to the trigger value, the IPI variance estimate given by Theorem 4 lies within the range of this sensitivity, which for realistic noise levels is small compared to the intrinsic variability of the variance across trials.

Chapter 14

Limitations

All Langevin models, including our proposed $14 \times 28\text{D}$ model, proceed from the assumption that the ion channel population is large enough (and the ion channel state transitions frequent enough) that the Gaussian approximations by which the white noise forcing terms are derived, are justified. Thus when the system size is too small, no Langevin system will be an appropriate. Fortunately the Langevin approximation appears quite accurate for realistic population sizes (cf. Tab. 3.1 and Figs. 11.4-11.5).

The $14 \times 28\text{D}$ model uses more noise sources than other approaches. However, stochastic shielding allows us to jettison noise sources that do not significantly impact the system dynamics (the voltage fluctuations and ISI distribution). Moreover, in order to compare the ISI distribution in detail among several variants of the Fox and Lu '94 model versus the Markov chain standard, we have considered a single value of the driving current, while other studies have compared parametrized responses such as the firing rate, ISI variance, or other moments, as a function of applied current. Accurate comparisons require large ensembles of independent trajectories, forcing a tradeoff between precision and breadth;

we opted here for precise comparisons at a representative level of the driving current.

From a conceptual point of view, a shortcoming of most Langevin models is the tendency for some channel state variables x to collide with the domain boundaries $x \in [0, 1]$ and to cross them during numerical simulations with finite time steps. We adopted the approach advocated by [90] of using “free boundaries” in which gating variables can make excursions into the (unphysical) range $x < 0$ or $x > 1$. Practically speaking, these excursions are always short, if the time step is reasonably small, as they tend to be self-correcting.¹ Another approach is to construct reflecting boundary conditions; different implementations of this idea were used in [20],[39] and [103]. Dangerfield’s method proved both slower and less accurate than the free boundary method. As an alternative method, one uses a biased rejection sampling approach, testing each gating variable of the 14D model on each time step, and repeating the noise sample for any time step violating the domain conditions [39, 103]. We found that this method had accuracy similar to that of Dangerfield’s method (L_1 -Wasserstein difference $\approx 4.4\text{e-}1$ msec, cf. Tab. 8.1) and runtime similar to that of the Fox and Lu 94 implementation, about 4 times slower than our 14D Langevin model.

Table 8.1 gives the accuracies with which each model reproduces the ISI distribution, compared to a standard reference distribution generated through a large number of samples of the MC method. The mean L_1 difference between a single sample and the reference sample is about 0.227 microseconds. For a nonnegative random variable $T \geq 0$, the difference in the mean under two probability distributions is bounded above by the L_1 difference in their cumulative distribution functions.² Thus the L_1 norm gives an idea of

¹To avoid complex entries, we use $|x|$ when calculating entries in the noise coefficient matrix.

²For a nonnegative random variable T with cumulative distribution function $F(t) = \mathbb{P}[T \leq t]$, the mean satisfies $E[T] = \int_0^\infty (1 - F(t)) dt$ [53]. Therefore the difference in mean under two distributions F_1 and F_2 satisfies $|E_1[T] - E_2[T]| = \left| \int_0^\infty F_1(t) - F_2(t) dt \right| \leq \rho_1(F_1, F_2)$.

the temporal accuracy with which one can approximate a given distribution by another. The mean difference between the ISI distribution generated by a single run of the full $14 \times 28\text{D}$ model is about $49 \mu\text{sec}$, and the discrepancy produced by the (significantly faster) SS model is about $76 \mu\text{sec}$. When would this level of accuracy matter for the function of a neuron within a network? The barn owl *Tyto alba* uses interaural time difference to localize its prey to within 1-2 degrees, a feat that requires encoding information in the precise timing of auditory system action potentials at the scale of 5-20 microseconds [45, 83]. For detailed studies of the effects of channel noise in this system, the superior accuracy of the MC model might be preferred. On the other hand, the timescale of information encoding in the human auditory nerve is thought to be in the millisecond range [50], with precision in the feline auditory system reported as low as $100 \mu\text{s}$ [62] (see also [132]). For these and other mammalian systems, the stochastic shielding approximation should provide sufficient accuracy.

Another limitation of our approach is that the interphase interval variance, σ_{IPI}^2 , is only well defined for the small noise region, and becomes prohibitively hard to calculate for large noise levels. Additionally, even given a stochastic trajectory and the associated Langevin model, σ_{IPI}^2 is computationally expensive to obtain. Fortunately, the theory we provided in this thesis give a good approximation to the decomposition of variance of ISIs up to a moderate noise level (cf. Figs. 11.4-11.5). Moreover, when we combine the stochastic shielding method with our $14 \times 28\text{D}$ Langevin model, the reduced model gives a good approximation to the full noise model for ion channel populations spanning the range of $[40, 10^5]$.

Like other approaches in the literature, our calculations are based on a linear approximation to the effects of the noise. However, Ito's formula (9.29) includes terms both of order $\sqrt{\epsilon}$ and ϵ . The latter weights the Hessian matrix of the asymptotic phase function,

$\partial_{ij}^2 \phi$. In our main result (10.8) we neglected the contribution of this higher order term. Similar truncations of either Taylor's expansion or Ito's formula are seen throughout the literature, for example eqn. (3.2.8) in [70], eqn. (120) in [110], and eqn. (2) in [31]. These authors favor an immediate phase reduction when ϵ is small, setting $\mathbf{X}(t) \approx \mathbf{X}_0(\theta(t))$ and

$$\frac{d\phi}{dt} = 1 + \sqrt{\epsilon} (\nabla \phi(\mathbf{X}))^\top \cdot \mathcal{G}(\mathbf{X}) \cdot d\mathbf{W}(t) \quad (14.1)$$

([110, 70]). As in Kuramoto's original phase reduction approach [70], we also evaluate the infinitesimal phase response curve \mathbf{Z} on the limit cycle throughout this thesis. This omission of the Hessian term could possibly cause additional discrepancies. Recent advances in the theory of nonlinear oscillators have provided means to obtain the asymptotic phase Hessian [3, 127, 129] but we have not attempted to implement these calculations for our 14D HH model.

Appendix A

Table of Common Symbols and Notations

Symbol	Meaning
C	Membrane capacitance ($\mu F/cm^2$)
v	Membrane potential (mV)
V_{Na}, V_K, V_L	Ionic reversal potential for Na^+ , K^+ and leak (mV)
\bar{g}_{ion}	maximal conductance for ion $\in \{Na^+, K^+\}$
I_{app}	Applied current to the membrane (nA/cm^2)
m, h, n	Dimensionless gating variables for Na^+ and K^+ channels
$\alpha_x, \beta_x, x \in \{m, n, h\}$	Voltage dependent rate constant ($1/msec$)
\mathbf{x}	vector of state variables
$\mathbf{M} = [M_1, M_2, \dots, M_8]$	Eight-component state vector for the Na^+ gates
$[m_{00}, \dots, m_{30}, m_{01}, \dots, m_{31}]^T$	Components for the Na^+ gates
$\mathbf{N} = [N_1, N_2, \dots, N_5]$	Five-component state vector for the K^+ gates
$[n_0, n_1, n_2, n_3, n_4]^T$	Components for the K^+ gates
M_{tot}, N_{tot}	Total number of Na^+ and K^+ channels

Table A.1: Common symbols and notations in this thesis (part I).

Symbol	Meaning
\mathcal{X}	4-dimensional manifold domain for 4D HH model
\mathcal{Y}	14-dimensional manifold domain for 14D HH model
Δ^k	k -dimensional simplex in \mathbb{R}^{k+1} given by $y_1 + \dots + y_{k+1} = 1, y_i \geq 0$
\mathcal{M}	Multinomial submanifold within the 14D space
$A_{\text{Na}}, A_{\text{K}}$	State-dependent rate matrix
D	State diffusion matrix
$\mathcal{G}, S, S_1, S_2, S_{\text{Na}}, S_{\text{K}}$	Noise coefficient matrices
ξ	Vector of independent δ -correlated Gaussian white noise with zero mean and unit variance
$\mathbf{X} = [X_1, X_2, \dots, X_d]$	A d -dimensional random variable for sample path
$\mathbf{W} = [W_1, W_2, \dots, W_n]$	A Wiener trajectory with n components
$\delta(\cdot)$	The Dirac delta function
δ_{ij}	The Kronecker delta
F_n	Empirical cumulative distribution function with n observations (in §8, we use m, n as sample sizes)
$e_i^{\text{Na}} \& e_i^{\text{K}}$	i th standard unit vector in $\mathbb{R}^8 \& \mathbb{R}^5$
$\zeta_k^{\text{ion}} = e_{j(k)}^{\text{ion}} - e_{i(k)}^{\text{ion}}$	stoichiometry vector for the k th edge, for ion $\in \{\text{Na}^+, \text{K}^+\}$
$\alpha_k(v)$	voltage-dependent per capita transition rate along k th edge
$i(k) \& j(k)$	source & destination nodes for k th edge
$M_{i(k)}$	fractional occupancy of source node for k th edge
$\mathbf{F}(\mathbf{X}) \& f(x)$	deterministic part of the evolution equation (mean-field)
\mathcal{D}	domain of the (stochastic) differential equation
\mathcal{Y}^0	“nullcline” surface associated with the voltage variable, where $f(v, \mathbf{m}, \mathbf{n}) = 0$
\mathcal{S}	arbitrary section transverse to the deterministic limit cycle

Table A.2: Common symbols and notations in this thesis (part II).

Symbol	Meaning
\mathcal{S}^u	isovoltage Poincaré section (where voltage is a constant u)
$\mathcal{S}_0^u, \mathcal{S}_+^u, \& \mathcal{S}_-^u$	“null”, “inward current” & “outward current” surface for voltage u and $f(v, \mathbf{m}, \mathbf{n}) = 0, f(v, \mathbf{m}, \mathbf{n}) > 0 \& f(v, \mathbf{m}, \mathbf{n}) < 0$, resp.
$\tau(\mathbf{x}, \mathcal{S})$	first passage time (FPT) from a point $\mathbf{x} \in \mathcal{D}$ to section \mathcal{S}
$T(\mathbf{x}, \mathcal{S})$	mean first passage time (MFPT) from point $\mathbf{x} \in \mathcal{D}$ to set \mathcal{S}
$S(\mathbf{x}, \mathcal{S})$	the second moment of the FPT from a point $\mathbf{x} \in \mathcal{D}$ to section \mathcal{S}
$\tau_k^u \& \tau_k^d$	k th voltage surface upcrossing & downcrossing time
I_k	k th interspike interval (ISI), for some reference voltage v_0
$I, H \& \sigma_{\text{ISI}}^2$	mean, 2nd moment, and variance of ISI
μ_k	k th iso-phase crossing time
Δ_k	k th iso-phase interval (IPI), for some reference phase ϕ_0
$\bar{T}_\epsilon, S_\epsilon, \sigma_{\text{IPI}}^2$	mean, 2nd moment, variance of iso-phase interval (for noise level ϵ)
$\sigma_{\phi,k}^2 \& \sigma_{\text{ISI},k}^2$	contribution of k th edge to the IPI variance and the ISI variance, resp.
$\gamma(t)$	deterministic limit cycle trajectory
\bar{T}_0	period of deterministic limit cycle
$\phi(\mathbf{x})$	asymptotic phase function for deterministic limit cycle
$\mathbf{Z}(t) = \nabla\phi(\gamma(t))$	infinitesimal phase response curve for deterministic limit cycle
\bar{T}_ϵ	mean period for noise level set to ϵ
$\bar{T}_1 = \left. \frac{\partial \bar{T}_\epsilon}{\partial \epsilon} \right _{\epsilon=0}$	sensitivity of the mean period to increasing noise level, in the small-noise limit
$T_\epsilon(\mathbf{x})$	mean–return-time (MRT) phase function for noise level set to ϵ
$T_1(\mathbf{x}) = \left. \frac{\partial T_\epsilon(\mathbf{x})}{\partial \epsilon} \right _{\epsilon=0}$	sensitivity of the phase function to noise in the small-noise limit
$T_0(\mathbf{x})$	MRT phase function for $\epsilon = 0$. Note $T_0(\mathbf{x}) = \text{const} - \bar{T}_0 \frac{\phi(\mathbf{x})}{2\pi}$ for an arbitrary constant

Table A.3: Common symbols and notations in this thesis (part III).

Appendix B

Parameters and Transition Matrices

Symbol	Meaning	Value
C	Membrane capacitance	$1 \mu F/cm^2$
\bar{g}_{Na}	Maximal sodium conductance	$120 \mu S/cm^2$
\bar{g}_K	Maximal potassium conductance	$36 \mu S/cm^2$
g_{leak}	Leak conductance	$0.3 \mu S/cm^2$
V_{Na}	Sodium reversal potential for Na^+	$50 mV$
V_K	Potassium reversal potential for K^+	$-77 mV$
V_{leak}	Leak reversal potential	$-54.4 mV$
I_{app}	Applied current to the membrane	$10 nA/cm^2$
\mathcal{A}	Membrane Area	$100 \mu m^2$
M_{tot}	Total number of Na^+ channels	6,000
N_{tot}	Total number of K^+ channels	18,00

Table B.1: Parameters used for simulations in this thesis.

Subunit kinetics for the Hodgkin and Huxley equations are given by

$$\alpha_m(v) = \frac{0.1(25 - v)}{\exp(2.5 - 0.1v) - 1} \quad (\text{B.1})$$

$$\beta_m(v) = 4 \exp(-v/18) \quad (\text{B.2})$$

$$\alpha_h(v) = 0.07 \exp(-v/20) \quad (\text{B.3})$$

$$\beta_h(v) = \frac{1}{\exp(3 - 0.1v) + 1} \quad (\text{B.4})$$

$$\alpha_n(v) = \frac{0.01(10 - v)}{\exp(1 - 0.1v) - 1} \quad (\text{B.5})$$

$$\beta_n(v) = 0.125 \exp(-v/80) \quad (\text{B.6})$$

$$A_{\mathbf{K}}(v) = \begin{bmatrix} D_{\mathbf{K}}(1) & \beta_n(v) & 0 & 0 & 0 \\ 4\alpha_n(v) & D_{\mathbf{K}}(2) & 2\beta_n(v) & 0 & 0 \\ 0 & 3\alpha_n(v) & D_{\mathbf{K}}(3) & 3\beta_n(v) & 0 \\ 0 & 0 & 2\alpha_n(v) & D_{\mathbf{K}}(4) & 4\beta_n(v) \\ 0 & 0 & 0 & \alpha_n(v) & D_{\mathbf{K}}(5) \end{bmatrix},$$

$$A_{\text{Na}} = \begin{bmatrix} D_{\text{Na}}(1) & \beta_m & 0 & 0 & \beta_h & 0 & 0 & 0 \\ 3\alpha_m & D_{\text{Na}}(2) & 2\beta_m & 0 & 0 & \beta_h & 0 & 0 \\ 0 & 2\alpha_m & D_{\text{Na}}(3) & 3\beta_m & 0 & 0 & \beta_h & 0 \\ 0 & 0 & \alpha_m & D_{\text{Na}}(4) & 0 & 0 & 0 & \beta_h \\ \alpha_h & 0 & 0 & 0 & D_{\text{Na}}(5) & \beta_m & 0 & 0 \\ 0 & \alpha_h & 0 & 0 & 3\alpha_m & D_{\text{Na}}(6) & 2\beta_m & 0 \\ 0 & 0 & \alpha_h & 0 & 0 & 2\alpha_m & D_{\text{Na}}(7) & 3\beta_m \\ 0 & 0 & 0 & \alpha_h & 0 & 0 & \alpha_m & D_{\text{Na}}(8) \end{bmatrix},$$

where the diagonal elements

$$D_{\text{ion}}(i) = - \sum_{j:j \neq i} A_{\text{ion}}(j, i), \quad \text{ion} \in \{\text{Na}, \text{K}\}.$$

Appendix C

Diffusion Matrix of the 14D Model

Define the state vector for Na⁺ and K⁺ channels as

$$\mathbf{M} = [m_{00}, m_{10}, m_{20}, m_{30}, m_{01}, m_{11}, m_{21}, m_{31}]^T,$$

and $\mathbf{N} = [n_0, n_1, n_2, n_3, n_4]^T$, respectively. The diffusion matrices D_{Na} and D_{K} are given by

$$D_{\text{Na}}^{(1:3)} = \begin{bmatrix} D_{\text{Na}}(1, 1) & -3\alpha_m m_{00} - \beta_m m_{10} & 0 \\ -3\alpha_m m_{00} - \beta_m m_{10} & D_{\text{Na}}(2, 2) & -2\alpha_m m_{10} - 2\beta_m m_{20} \\ 0 & -2\alpha_m m_{10} - 2\beta_m m_{20} & D_{\text{Na}}(3, 3) \\ 0 & 0 & -\alpha_m m_{20} - 3\beta_m m_{30} \\ -\alpha_h m_{00} - \beta_h m_{01} & 0 & 0 \\ 0 & -\alpha_h m_{10} - \beta_h m_{11} & 0 \\ 0 & 0 & -\alpha_h m_{20} - \beta_h m_{21} \\ 0 & 0 & 0 \end{bmatrix},$$

$$D_{\text{Na}}^{(4:6)} = \begin{bmatrix} 0 & -\alpha_h m_{00} - \beta_h m_{01} & 0 \\ 0 & 0 & -\alpha_h m_{10} - \beta_h m_{11} \\ -\alpha_m m_{20} - 3\beta_m m_{30} & 0 & 0 \\ D_{\text{Na}}(4, 4) & 0 & 0 \\ 0 & D_{\text{Na}}(5, 5) & -3\alpha_m m_{01} - \beta_m m_{11} \\ 0 & -3\alpha_m m_{01} - \beta_m m_{11} & D_{\text{Na}}(6, 6) \\ 0 & 0 & -2\alpha_m m_{11} - 2\beta_m m_{21} \\ -\alpha_h m_{30} - \beta_h m_{31} & 0 & 0 \end{bmatrix},$$

$$D_{\text{Na}}^{(7:8)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -\alpha_h m_{20} - \beta_h m_{21} & 0 \\ 0 & -\alpha_h m_{30} - \beta_h m_{31} \\ 0 & 0 \\ -2\alpha_m m_{11} - 2\beta_m m_{21} & 0 \\ D_{\text{Na}}(7, 7) & -\alpha_m m_{21} - 3\beta_m m_{31} \\ -\alpha_m m_{21} - 3\beta_m m_{31} & D_{\text{Na}}(8, 8) \end{bmatrix},$$

where

$$D_{\text{ion}}(i, i) = - \sum_{j: j \neq i} D_{\text{ion}}(j, i), \text{ for ion} \in \{\text{Na}, \text{K}\}.$$

Bibliography

- [1] L.F Abbott. Lopicque’s introduction of the integrate-and-fire model neuron (1907). *Brain Research Bulletin*, 50(5–6):303 – 304, 1999.
- [2] Edward J Allen, Linda JS Allen, Armando Arciniega, and Priscilla E Greenwood. Construction of equivalent stochastic differential equation models. *Stochastic Analysis and Applications*, 26(2):274–297, 2008.
- [3] Z. Aminzare, P. Holmes, and V. Srivastava. On phase reduction and time period of noisy oscillators. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 4717–4722, 2019.
- [4] 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.
- [5] David F. Anderson, Arnab Ganguly, and Thomas G. Kurtz. Error analysis of tau-leap simulation methods. *Ann. Appl. Probab.*, 21(6):2226–2262, December 2011.
- [6] David F Anderson and Thomas G Kurtz. *Stochastic Analysis of Biochemical Systems*, volume 1. Springer, 2015.
- [7] Rony Azouz and Charles M Gray. Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo. *Proceedings of the National Academy of Sciences*, 97(14):8110–8115, 2000.
- [8] Mark W Barnett and Philip M Larkman. The action potential. *Practical Neurology*, 7(3):192–197, 2007.
- [9] Francisco Bezanilla. Single sodium channels from the squid giant axon. *Biophysical Journal*, 52(6):1087–1090, 1987.
- [10] Paul C Bressloff and Jay M Newby. Path integrals and large deviations in stochastic hybrid systems. *Physical Review E*, 89(4):042701, 2014.
- [11] Eric Brown, Jeff Moehlis, and Philip Holmes. On the phase reduction and response dynamics of neural oscillator populations. *Neural Comput*, 16(4):673–715, Apr 2004.
- [12] Ian C Bruce. Evaluation of stochastic differential equation approximation of ion channel gating models. *Annals of Biomedical Engineering*, 37(4):824–838, 2009.
- [13] Nicolas Brunel and Peter E Latham. Firing rate of the noisy quadratic integrate-and-fire neuron. *Neural Computation*, 15(10):2281–2306, 2003.

- [14] Evelyn Buckwar and Martin G Riedler. An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *Journal of Mathematical Biology*, 63(6):1051–1093, 2011.
- [15] Robert J Butera Jr, John Rinzel, and Jeffrey C Smith. Models of respiratory rhythm generation in the pre-botzinger complex. i. bursting pacemaker neurons. *Journal of Neurophysiology*, 82(1):382–397, 1999.
- [16] Daniela Calvetti and Erkki Somersalo. *Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*. Springer, 2007.
- [17] Brett C Carter, Andrew J Giessel, Bernardo L Sabatini, and Bruce P Bean. Transient sodium current at subthreshold voltages: activation by epsp waveforms. *Neuron*, 75(6):1081–93, Sep 2012.
- [18] J R Clay and L J DeFelice. Relationship between membrane excitability and single channel open-close kinetics. *Biophys J*, 42(2):151–7, May 1983.
- [19] David Colquhoun and AG Hawkes. On the stochastic properties of single ion channels. *Proc. R. Soc. Lond. B*, 211(1183):205–235, 1981.
- [20] CE Dangerfield, David Kay, and Kevin Burrage. Stochastic models and simulation of ion channel dynamics. *Procedia Computer Science*, 1(1):1587–1596, 2010.
- [21] Ciara E Dangerfield, David Kay, and Kevin Burrage. Modeling ion channel dynamics through reflected stochastic differential equations. *Physical Review E*, 85(5):051907, 2012.
- [22] Peter Dayan and L. F. Abbott. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Computational Neuroscience. The MIT Press, 2001.
- [23] Erik De Schutter and James M Bower. An active membrane model of the cerebellar Purkinje cell I. Simulation of current clamps in slice. *Journal of neurophysiology*, 71(1):375–400, 1994.
- [24] Erik De Schutter and James M Bower. An active membrane model of the cerebellar Purkinje cell II. Simulation of synaptic responses. *Journal of Neurophysiology*, 71(1):401–419, 1994.
- [25] Erik De Schutter and James M Bower. Simulated responses of cerebellar Purkinje cells are independent of the dendritic location of granule cell synaptic inputs. *Proceedings of the National Academy of Sciences*, 91(11):4736–4740, 1994.

- [26] Alan D Dorval and John A White. Channel noise is essential for perithreshold oscillations in entorhinal stellate neurons. *Journal of neuroscience*, 25(43):10025–10028, 2005.
- [27] JE Dowling. An introduction to neuroscience, 1992.
- [28] Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, 27(3):642–669, 1956.
- [29] B. Earnshaw and J. Keener. Global asymptotic stability of solutions of nonautonomous master equations. *SIAM Journal on Applied Dynamical Systems*, 9(1):220–237, 2010.
- [30] B. Earnshaw and J. Keener. Invariant manifolds of binomial-like nonautonomous master equations. *SIAM Journal on Applied Dynamical Systems*, 9(2):568–588, 2010.
- [31] G Bard Ermentrout, Bryce Beverlin, 2nd, Todd Troyer, and Theoden I Netoff. The variance of phase-resetting curves. *J Comput Neurosci*, Jan 2011.
- [32] G. Bard Ermentrout and David H. Terman. *Foundations Of Mathematical Neuroscience*. Springer, 2010.
- [33] A Aldo Faisal and Simon B Laughlin. Stochastic simulations on the reliability of action potential propagation in thin axons. *PLoS Computational Biology*, 3(5), 2007.
- [34] A Aldo Faisal, John A White, and Simon B Laughlin. Ion-channel noise places limits on the miniaturization of the brain’s wiring. *Current Biology*, 15(12):1143–1149, 2005.
- [35] Elizabeth M Fenwick, A Marty, and E Neher. Sodium and calcium channels in bovine chromaffin cells. *The Journal of physiology*, 331(1):599–635, 1982.
- [36] Fernando R Fernandez, Jordan D T Engbers, and Ray W Turner. Firing dynamics of cerebellar Purkinje cells. *J Neurophysiol*, 98(1):278–94, Jul 2007.
- [37] Michael D Forrest. Intracellular calcium dynamics permit a purkinje neuron model to perform toggle and gain computations upon its inputs. *Frontiers in computational neuroscience*, 8:86, 2014.
- [38] Michael D Forrest. Simulation of alcohol action upon a detailed Purkinje neuron model and a simpler surrogate model that runs >400 times faster. *BMC Neuroscience*, 16(1):27, 2015.

- [39] Fox and Lu. Emergent collective behavior in large numbers of globally coupled independently stochastic ion channels. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 49(4):3421–3431, Apr 1994.
- [40] Ronald F Fox. Stochastic versions of the Hodgkin-Huxley equations. *Biophysical Journal*, 72(5):2068–2074, 1997.
- [41] Ronald F Fox. Critique of the Fox-Lu model for Hodgkin-Huxley fluctuations in neuron ion channels. *arXiv preprint arXiv:1807.07632*, 2018.
- [42] Chetan Gadgil, Chang Hyeong Lee, and Hans G. Othmer. A stochastic analysis of first-order reaction networks. *Bulletin of Mathematical Biology*, 67, 2005.
- [43] Crispin W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer Verlag, 2nd edition, 2004.
- [44] George L Gerstein. Analysis of firing patterns in single neurons. *Science*, 131(3416):1811–1812, 1960.
- [45] Wulfram Gerstner, Richard Kempter, J Leo Van Hemmen, and Hermann Wagner. A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383(6595):76–78, 1996.
- [46] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81:2340–2361, 1977.
- [47] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.
- [48] 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.
- [49] Joshua H. Goldwyn and Eric Shea-Brown. The what and where of adding channel noise to the Hodgkin-Huxley equations. *PLoS Comput Biol*, 7(11):e1002247, November 2011.
- [50] Joshua H Goldwyn, Eric Shea-Brown, and Jay T Rubinstein. Encoding and decoding amplitude-modulated cochlear implant stimuli – a point process analysis. *Journal of Computational Neuroscience*, 28(3):405–424, 2010.
- [51] Jorge Golowasch, Gladis Thomas, Adam L Taylor, Arif Patel, Arlene Pineda, Christopher Khalil, and Farzan Nadim. Membrane capacitance measurements revisited: dependence of capacitance value on measurement method in nonisopotential neurons. *Journal of neurophysiology*, 102(4):2161–2175, 2009.

- [52] Priscilla E Greenwood, Mark D McDonnell, and Lawrence M Ward. Dynamics of gamma bursts in local field potentials. *Neural Computation*, 27(1):74–103, 2015.
- [53] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. Oxford University Press, third edition, 2005.
- [54] J. Guckenheimer. Isochrons and phaseless sets. *Journal of Mathematical Biology*, 1:259–273, 1975. 10.1007/BF01273747.
- [55] Marifi Güler. An investigation of the stochastic Hodgkin-Huxley models under noisy rate functions. *Neural Computation*, 25(9):2355–2372, 2013.
- [56] Marifi Güler. Stochastic Hodgkin-Huxley equations with colored noise terms in the conductances. *Neural Computation*, 25(1):46–74, 2013.
- [57] Boris S Gutkin and G Bard Ermentrout. Dynamics of membrane excitability determine interspike interval variability: a link between spike generation mechanisms and cortical spike train statistics. *Neural Computation*, 10(5):1047–1065, 1998.
- [58] Terrell L Hill and Yi-Der Chen. On the theory of ion transport across the nerve membrane: IV. noise from the open-close kinetics of K⁺ channels. *Biophysical Journal*, 12(8):948–959, 1972.
- [59] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117:500–544, 1952.
- [60] Yandong Huang, Sten Rüdiger, and Jianwei Shuai. Channel-based Langevin approach for the stochastic Hodgkin-Huxley neuron. *Physical Review E*, 87(1):012716, 2013.
- [61] Yandong Huang, Sten Rüdiger, and Jianwei Shuai. Accurate Langevin approaches to simulate Markovian channel dynamics. *Physical Biology*, 12(6):061001, 2015.
- [62] Nikita S Imenov and Jay T Rubinstein. Stochastic population model for electrical stimulation of the auditory nerve. *IEEE Transactions on Biomedical Engineering*, 56(10):2493–2501, 2009.
- [63] James P Keener. Invariant manifold reductions for Markovian ion channel dynamics. *J Math Biol*, 58(3):447–57, Mar 2009.
- [64] James P Keener. Exact reductions of Markovian dynamics for ion channels with a single permissive state. *J Math Biol*, 60(4):473–9, Apr 2010.

- [65] James P Keener and James Sneyd. *Mathematical Physiology*, volume 1. Springer, 1998.
- [66] Zayd M Khaliq, Nathan W Gouwens, and Indira M Raman. The contribution of resurgent sodium current to high-frequency firing in Purkinje neurons: an experimental and modeling study. *The Journal of Neuroscience*, 23(12):4899–4912, 2003.
- [67] Tilman Kispersky and John A White. Stochastic models of ion channel gating. *Scholarpedia*, 3(1):1327, 2008.
- [68] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Number 23 in Applications of Mathematics. Springer, 1999.
- [69] Andrey Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Inst. Ital. Attuari, Giorn.*, 4:83–91, 1933.
- [70] Yoshiki Kuramoto. *Chemical Oscillations Waves and Turbulence*. Dover, 2003.
- [71] Thomas G. Kurtz. *Approximation of population processes*. Society for Industrial and Applied Mathematics, 1981.
- [72] Andrzej Lasota and Michael C. Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, volume 97 of *Applied Mathematical Sciences*. Springer-Verlag New York, 1994.
- [73] 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.
- [74] Benjamin Lindner. Interspike interval statistics of neurons driven by colored noise. *Physical Review E*, 69(2):022901, 2004.
- [75] Benjamin Lindner, Alexander Cao, and Peter Thomas. A partial differential equation for the mean first-return-time phase of a stochastic oscillator. In *APS Meeting Abstracts*, 2019.
- [76] Rodolfo Llinás and M Sugimori. Electrophysiological properties of in vitro Purkinje cell dendrites in mammalian cerebellar slices. *The Journal of physiology*, 305(1):197–213, 1980.
- [77] Rodolfo Llinás and M Sugimori. Electrophysiological properties of in vitro Purkinje cell somata in mammalian cerebellar slices. *The Journal of physiology*, 305(1):171–195, 1980.

- [78] Zachary F Mainen and Terrence J Sejnowski. Reliability of spike timing in neocortical neurons. *Science*, 268(5216):1503–1506, 1995.
- [79] Jonathan C Makielski, Michael F Sheets, Dorothy A Hanck, Craig T January, and Harry A Fozzard. Sodium current in voltage clamped internally perfused canine cardiac Purkinje cells. *Biophysical Journal*, 52(1):1–11, 1987.
- [80] Alan J McKane and Timothy J Newman. Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical review letters*, 94(21):218102, 2005.
- [81] Hiroyuki Mino, Jay T Rubinstein, and John A White. Comparison of algorithms for the simulation of action potentials with stochastic sodium channels. *Ann Biomed Eng*, 30(4):578–587, Apr 2002.
- [82] Tsugumichi Miyasho, Hiroshi Takagi, Hideo Suzuki, Shigeo Watanabe, Masashi Inoue, Yoshihisa Kudo, and Hiroyoshi Miyakawa. Low-threshold potassium channels and a low-threshold calcium channel regulate Ca^{2+} spike firing in the dendrites of cerebellar Purkinje neurons: a modeling study. *Brain Research*, 891(1):106–115, 2001.
- [83] Andrew Moiseff and Masakazu Konishi. The owl’s interaural pathway is not involved in sound localization. *Journal of Comparative Physiology*, 144(3):299–304, 1981.
- [84] LM Mukhametov, G Rizzolatti, and V Tradardi. Spontaneous activity of neurones of nucleus reticularis thalami in freely moving cats. *The Journal of physiology*, 210(3):651–667, 1970.
- [85] Erwin Neher and Bert Sakmann. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature*, 260(5554):799, 1976.
- [86] Theoden Netoff, Michael A Schwemmer, and Timothy J Lewis. Experimentally estimating phase response curves of neurons: theoretical and practical issues. In *Phase Response Curves in Neuroscience*, pages 95–129. Springer, 2012.
- [87] Jay M Newby, Paul C Bressloff, and James P Keener. Breakdown of fast-slow analysis in an excitable system with channel noise. *Phys Rev Lett*, 111(12):128101, Sep 2013.
- [88] Cian O’Donnell and Mark CW van Rossum. Spontaneous action potentials and neural coding in unmyelinated axons. *Neural Computation*, 27(4):801–818, 2015.
- [89] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 6th edition, 2007.

- [90] Patricio Orio and Daniel Soudry. Simple, fast and accurate implementation of the diffusion approximation algorithm for stochastic ion channels with multiple states. *PLoS One*, 7(5):e36670, 2012.
- [91] Saak V Ovsepiyan and David D Friel. The leaner P/Q-type calcium channel mutation renders cerebellar Purkinje neurons hyper-excitable and eliminates Ca^{2+} - Na^{+} spike bursts. *European Journal of Neuroscience*, 27(1):93–103, 2008.
- [92] Khashayar Pakdaman, Michele Thieullen, and Gilles Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Advances in Applied Probability*, 42(3):761–794, 2010.
- [93] Danilo Pezo, Daniel Soudry, and Patricio Orio. Diffusion approximation-based simulation of stochastic ion channels: which method to use? *Frontiers in Computational Neuroscience*, 8:139, 2014.
- [94] Daniela Pietrobon. Ca^{2+} channelopathies. *Pflügers Archiv-European Journal of Physiology*, 460(2):375–393, 2010.
- [95] Shusen Pu and Peter J. Thomas. Fast and accurate Langevin simulations of stochastic Hodgkin-Huxley dynamics. *Neural Computation*, 32(10):1775–1835, 2020.
- [96] Shusen Pu and Peter J. Thomas. Resolving molecular contributions of ion channel noise to interspike interval variability through stochastic shielding. In preparation, 2020.
- [97] Sanjeev Rajakulendran, Diego Kaski, and Michael G Hanna. Neuronal P/Q-type calcium channel dysfunction in inherited disorders of the CNS. *Nature Reviews Neurology*, 8(2):86, 2012.
- [98] I M Raman and B P Bean. Inactivation and recovery of sodium currents in cerebellar Purkinje neurons: evidence for two mechanisms. *Biophys J*, 80(2):729–37, Feb 2001.
- [99] Hannes Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer, 2nd edition, 1996.
- [100] Peter Rowat. Interspike interval statistics in the stochastic Hodgkin-Huxley model: Coexistence of gamma frequency bursts and highly irregular firing. *Neural Computation*, 19(5):1215–1250, 2007.
- [101] Peter F Rowat and Priscilla E Greenwood. Identification and continuity of the distributions of burst-length and interspike intervals in the stochastic Morris-Lecar neuron. *Neural computation*, 23(12):3094–3124, 2011.

- [102] Nicolaus T Schmandt and Roberto F Galán. Stochastic-shielding approximation of Markov chains and its application to efficiently simulate random ion-channel gating. *Physical Review Letters*, 109(11):118101, 2012.
- [103] Gerhard Schmid, Igor Goychuk, and Peter Hänggi. Stochastic resonance as a collective property of ion channel assemblies. *EPL (Europhysics Letters)*, 56(1):22, 2001.
- [104] Deena R Schmidt, Roberto F Galán, and Peter J Thomas. Stochastic shielding and edge importance for Markov chains with timescale separation. *PLoS Computational Biology*, 14(6):e1006206, 2018.
- [105] Deena R Schmidt and Peter J Thomas. Measuring edge importance: a quantitative analysis of the stochastic shielding approximation for random processes on graphs. *The Journal of Mathematical Neuroscience*, 4(1):6, 2014.
- [106] Otto H Schmitt. A thermionic trigger. *Journal of Scientific instruments*, 15(1):24, 1938.
- [107] 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.
- [108] J. Schwabedal and A. Pikovsky. Effective phase dynamics of noise-induced oscillations in excitable systems. *Phys Rev E Stat Nonlin Soft Matter Phys*, 81(4 Pt 2):046218, Apr 2010.
- [109] J. Schwabedal and A. Pikovsky. Phase description of stochastic oscillations. *Phys. Rev. Lett.*, 110:4102, 2013.
- [110] Michael A Schwemmer and Timothy J Lewis. The theory of weakly coupled oscillators. In *Phase response curves in neuroscience*, pages 3–31. Springer, 2012.
- [111] B Sengupta, SB Laughlin, and JE Niven. Comparison of Langevin and Markov channel noise models for neuronal signal generation. *Physical Review E*, 81(1):011918, 2010.
- [112] Michael N Shadlen and William T Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18(10):3870–3896, 1998.
- [113] Ryuuzo Shingai and Fred N Quandt. Single inward rectifier channels in horizontal cells. *Brain Research*, 369(1-2):65–74, 1986.

- [114] Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*. SIAM, 2009.
- [115] FJ Sigworth. Sodium channels in nerve apparently have two conductance states. *Nature*, 270(5634):265–267, 1977.
- [116] E Skaugen and L Walløe. Firing behaviour in a stochastic nerve membrane model based upon the Hodgkin-Huxley equations. *Acta Physiol Scand*, 107(4):343–63, Dec 1979.
- [117] Nickolay Smirnov. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19(2):279–281, 1948.
- [118] Richard B Stein. A theoretical analysis of neuronal variability. *Biophysical Journal*, 5(2):173–194, 1965.
- [119] 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.
- [120] Daniel T O’Connor, Sushil K Mahata, Manjula Mahata, Qijiao Jiang, Vivian Y Hook, and Laurent Taupenot. Primary culture of bovine chromaffin cells. *Nature protocols*, 2(5):1248–1253, 2007.
- [121] Peter J Thomas and Benjamin Lindner. Asymptotic phase for stochastic oscillators. *Physical review letters*, 113(25):254101, 2014.
- [122] Rafael D Vilela and Benjamin Lindner. Are the input parameters of white noise driven integrate and fire neurons uniquely determined by rate and CV? *Journal of Theoretical Biology*, 257(1):90–99, 2009.
- [123] 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.
- [124] Leonid N Wasserstein. Markov processes over denumerable products of spaces describing large systems of automata. *Problems of Information Transmission*, 5(3):47–52, 1969.
- [125] 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.
- [126] John A White, Jay T Rubinstein, and Alan R Kay. Channel noise in neurons. *Trends in Neurosciences*, 23(3):131–137, 2000.

- [127] Dan Wilson and Bard Ermentrout. Greater accuracy and broadened applicability of phase reduction using isostable coordinates. *Journal of mathematical biology*, 76(1-2):37–66, 2018.
- [128] Dan Wilson and Bard Ermentrout. An operational definition of phase characterizes the transient response of perturbed limit cycle oscillators. *SIAM Journal on Applied Dynamical Systems*, 17(4):2516–2543, 2018.
- [129] Dan Wilson and Bard Ermentrout. Augmented phase reduction of (not so) weakly perturbed coupled oscillators. *SIAM Review*, 61(2):277–315, 2019.
- [130] A. T. Winfree. Patterns of phase compromise in biological cycles. *Journal of Mathematical Biology*, 1:73–95, 1974.
- [131] A.T. Winfree. *The Geometry of Biological Time*. Springer-Verlag, New York, second edition, 2000.
- [132] J. Woo, C. A. Miller, and P. J. Abbas. Simulation of the electrically stimulated cochlear neuron: modeling adaptation to trains of electric pulses. *IEEE Transactions on Biomedical Engineering*, 56(5):1348–1359, 2009.
- [133] Santiago Ramón y Cajal. *Histologie du système nerveux de l’homme et des vertébrés*, volume 2. A. Maloine, 1911.
- [134] Haitao Yu, Rishi R Dhingra, Thomas E Dick, and Roberto F Galán. Effects of ion channel noise on neural circuits: an application to the respiratory pattern generator to investigate breathing variability. *Journal of Neurophysiology*, 117(1):230–242, 2017.