MODELS AND ANALYSIS OF A STOCHASTIC NEURAL SOURCE CODER

BY

ROBIN SINGH SIDHU

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Advisers:

Professor Douglas L. Jones
Adjunct Professor Rama Ratnam

ABSTRACT

Information transfer in neurons takes place through action potentials (spikes) which are metabolically expensive. A neural coding approach was developed by Johnson et al. (2016) that is optimal, high-fidelity, energy-efficient and well matches the experimental spiking behavior of real neurons. This coder, called a neural source-coder, uses an adaptive threshold to internally reconstruct the stimulus. The spikes are timed to minimize coding error. These spikes are generated by a deterministic firing rule. However, some random variability in spike timing is observed in real data. It seems reasonable to account for the variability by adding a stochastic component to the deterministic model.

Previously, the source-coding neuron used a constant threshold for generating spikes, while the stochastic neural encoder uses a partially randomized threshold. In this thesis we explore this random component and its success in explaining and recreating experimentally obtained spike-train statistics (from P-type electrosensory afferents of weakly electric fish). We also take a close look at the growth in variance of inter-spike intervals (ISIs) and the deviation of the stochastic source-coding neuron from an ideal DC-block system with infinite memory. The stochastic source-coding neuron model was able to achieve very accurate reconstructions of P-type weakly electric fish spike-time statistics (inter-spike interval histograms, serial correlation coefficients) with a very simple model consisting of only four free tuning parameters. We were also able to demonstrate a markedly slower growth in variance consistent with experimental data but which Poisson spike trains fail to capture. We were also able to derive mathematical correspondence for the observed experimental behavior such as the SCC trends, the rate of growth in variance of the ISIs and the power spectrum of the spike trains at low frequencies. The simulations back the mathematical findings illustrating the success of the model at creating statistically accurate and realistic spike trains.

To my advisers, Prof. D.L Jones and Prof. Ratnam, for helping me through tough times and their continued guidance. To my parents and sisters, for their love and support. To Blazer, for being Blazer.

ACKNOWLEDGMENTS

This research was financially supported by ADSC Incentive Funds allocated to Professor Douglas L. Jones. The electric fish data used in the thesis was collected by Professor Rama Ratnam, Zhian Xu, and Noura Sharabash in the laboratory of Professor Mark E. Nelson, UIUC. Professor Douglas L. Jones has helped me tremendously with the mathematical analysis of the SCCs, growth in variance, and spectral density of spike trains and general insights in the model simulations. I also benefited greatly from Professor Ratnam's insight into neuro-physiological mechanisms for a better mathematical modelling and assistance in statistical representation of the experimental data. My labmates Alex Asilador and Duc Phan have also been a great source of inspiration for this project.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Information transfer in biological systems is attributed to networks of neurons. Neural coding studies the information flow from the external stimulus to the brain's response. Since every physiologically vital decision is conveyed in our biology through neural coding, bettering our understanding of the neural code is paramount. Modelling neuron spikes allows us to quantitatively characterize the flow of information and how external stimulus is represented in action potentials (spikes). Neural coding comprises of encoding (which is the map from external stimulus and the neuronal response) and decoding (which is the reverse map from response to stimulus). Our stochastic source-coding neuron model is based on a deterministic model [1]. The stochastic source-coding neuron model (Figure 1.1) is similar to the deterministic one in that it is also made up of an encoder and decoder. The encoder generates spikes given a stimulus $s(t)$, and the decoder filters the spikes with a reconstruction filter $h(t)$ and generates an internal reconstruction $r(t)$. Johnson et al. [1] postulated that the decoder is the neuron's time-varying threshold, modeled as a low-pass filter to mimic a post-synaptic RC-membrane. The encoder tracks the coding error $s(t) - r(t)$ and fires a spike whenever the error reaches a stimulus-dependent firing threshold $\gamma(s, t)$. For a fixed spike-rate R, spikes are timed to minimize the mean-squared reconstruction (coding) error by optimization of the threshold. The difference between the deterministic and the stochastic encoder is that the former does not just have a deterministic threshold $\gamma(s)$ but also a stochastic component $\gamma(s) + x[i(t)]$ where i is an integer denoting the spike number (details in subsequent sections).

Following [1], we make the simple supposition that noise is added to the stimulus-dependent threshold $\gamma(s)$. To make the stochastic source-coding neuron realistic, we adjust the threshold noise statistics to best match the

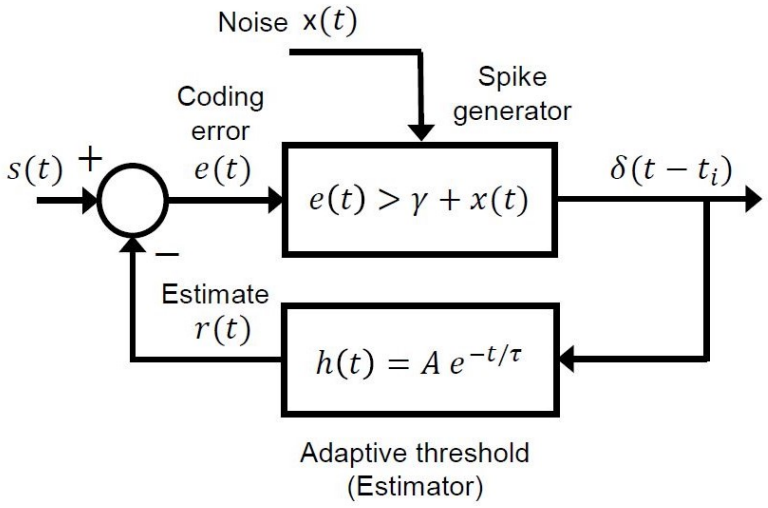

Figure 1.1: The stochastic source-coding neuron. The stimulus is represented by $s(t)$. $h(t)$ is the reconstruction filter. The error signal $e(t) = s(t) - r(t)$ is then compared to the stochastic threshold $\gamma + x(t)$, with spikes firing every time the threshold is crossed.

data observed from real neurons. While the most common assumption is that neuronal spike-times follow a Poisson process [2], here we consider non-renewal statistics, in particular anti-correlated ISIs [3], and our findings strongly argue against the renewal assumption (for electrosensory afferent spike trains at the very least). A dynamic or adaptive threshold was originally proposed as an empirical mechanism that generates negative serial correlations in the ISI sequence [4]. Anti-correlations also stabilize the mean spike firing rate and have been presumed to serve the ethologically important function of weak-signal detection [5, 3], particularly in sequential or real-time tasks [6, 7]. While sensory neurons in several model systems are known to exhibit such correlations (see [7, 8] for a review), the P-type primary electroreceptor afferents of a weakly electric fish demonstrate some of the strongest known anti-correlations [3] and have been successfully modeled using a dynamic threshold [9, 5, 10]. The proposed source-coding neuron was developed from observations of P-type spike trains [11], albeit deterministically, and has been shown to produce dynamic threshold behavior. It was therefore reasonable to expect (and consequently observed) that a stochastic extension of the source-coding neuron will also exhibit anti-correlations,

and it is of interest to ask how these correlations can be introduced into a spike-train model, and to determine their effects on spike-time statistics.

Thus, there is substantial evidence, both computational and experimental, that noise in neural systems serves to enhance coding and processing of sensory signals. Johnson [12] argues and shows by simulations that while the addition of stochastic spike firing will only increase encoding error for single neurons, it could lead to greatly reduced encoding error in a population of neurons.

In this thesis our primary focus is on a neural coding model that generates spike trains with all statistics closely resembling those of experimentally obtained ones while at the same time avoiding overly convoluted mathematical modelling with a plethora of exception handling and many parameters that usually lead to over-fitting. The strength and success of the model lie in the fact that it explains almost all statistically significant spike-train characteristics with a handful of model parameters. We briefly touch upon related work in the field in Chapter 2, which shall clearly establish the relevance and ingenuity of the model. Chapter 3 describes the model with a particular focus on the stochastic threshold. In Chapter 4 we take a step further from the simplifying assumptions in Chapter 3 and aim to encompass more detailed statistics and variance growth. Chapter 5 includes all the results and their comparison with the experimentally obtained data. Finally we conclude our findings in Chapter 6 and touch upon some future directions and applications for the model.

CHAPTER 2

RELATED WORK

Coding by sensory neurons is presumed to be constrained by metabolic energy [13, 14, 15, 16, 17, 18, 19], which places limits on the excitability of neurons. Thus, an energy constraint would demand that neurons put out as few spikes as possible. On the other hand, sensory processing requires high-fidelity coding in the transduction process in order to preserve information about critical stimulus features. Out of necessity, high-fidelity coding would demand a high spike rate. Thus there is an energy fidelity trade-off. In [11, 20] our group proposed a biophysically plausible coding framework suggesting that the trade-off is obtained by optimally timing spikes, so that coding fidelity is maximized for a given constraint on the firing rate of the neuron. In this framework, it was shown that the neuron's dynamic threshold is an internal reconstruction of the stimulus and maintains an ongoing estimate of the coding error. This neural coding mechanism is broadly similar to lossy source-coding in digital systems, where the coding error is minimized for a fixed bit rate [21].

Other work has investigated neural coding strategies which balance a trade-off between coding fidelity and expended energy. Some approaches aim to maximize spike-train entropy subject to a constraint on the energy or spike rate [22, 23], and these approaches can be used to study inter-spike interval codes [24]. An alternative method has been to characterize the neurons as communication channels which maximize a ratio of bits transmitted to energy expended [25, 26, 27]. This is a channel-coding approach, as opposed to a source-coding approach [28].

The source-coding neuron has some elements in common with previous approaches that employ adaptive or dynamic thresholds [29, 10, 9, 30] to generate spike trains that capture either the serial corrections in the inter-spike intervals (ISIs) or match spike times. In most previous work the integrator is reset after a spike is output [29, 10], and only a few previous reports do

not reset the integrator [9, 30]. None of these adaptive or dynamic threshold models incorporate an energy constraint or optimize coding fidelity. An alternative interpretation of a feedback signal is the noise-shaping neuron hypothesis proposed by Shin et al. [31, 32, 33]. The noise-shaping hypothesis proposes that the feedback filter is specifically designed to adapt the neuron's input/output function to the input signal statistics and filter out encoding noise from the signal band. The term noise-shaping is used in the same sense as noise-shaping in oversampled delta-sigma modulation [34]. While there are similarities between the noise-shaping neuron and the source-coding neuron, previous work on the noise-shaping neuron [32] is not concerned with coding fidelity or an energy constraint.

In [11, 20] the source-coding neuron was deterministic. Here we provide a stochastic extension, which is an important step in modelling neural systems. Physiological neurons exhibit some variability in their spike times even when stimulated repeatedly with a deterministic stimulus [35, 36]. On the other hand, the spikes generated by the source-coding neuron are deterministically timed so as to minimize coding error. Thus, for the same stimulus and a fixed spike-rate R, any jittered set of spike times will increase coding error.

When studying neuronal noise, a classical view is that neurons are simply inherently noisy due to the underlying biophysics, and that neuronal networks function despite these noisy mechanisms [37]. Over the last twenty years, however, many studies have investigated the enhancement of neurons and neural networks by the addition of noise. The idea of noise-enhanced processing is similar to the concept of dithering in quantization [38], and is often termed stochastic resonance [39, 40, 41]. Additive threshold noise has been shown to enhance processing in both single neuron models [42] and populations of neuron models [43, 44]. In particular, additive noise in a population of neuron models improves the signal-to-noise ratio of reconstructed waveforms when using a simple summing architecture [43]. Experimental evidence for stochastic resonance in neuronal systems has been observed in a range of systems such as cutaneous mammalian mechano-receptors [43] and cricket sensory systems [45].

Jones et al. [11], who introduced the deterministic source-coder, reported that the model added or deleted some spikes in comparison with experimental data. However, by and large, we noticed that spikes from the deterministic source-coding neuron were simply jittered in comparison to the measured

spike times. Here, we incorporate a mechanism that induces timing jitter in a source-coding neuron, and we study whether it has any benefits when representing sensory signals.

CHAPTER 3

A STOCHASTIC EXTENSION TO THE SOURCE-CODING NEURON

The deterministic source-coding neuron was able to explain many spike time characteristics such as the anti-correlations in spike trains (non-renewal), key features observed in peristimulus time histograms (PSTH), short-term adaptation, and a DC-block in the power spectral density of the spike train. A better model must account for the variance in spike trains, and reproduce the experimental ISI histogram and serial correlation coefficients (SCC), hence our motivation for a stochastic threshold extension of the source-coder.

3.1 Stochastic Source-Coding Neuron

In this section we introduce a stochastic threshold neuron model and obtain general expressions for the serial correlation coefficients for ISIs of the spike train generated by it. This model improves over [12] in that we now have mathematical justification for the choice of the noisy threshold spectrum. We were also able to quantitatively identify how the model parameters affected the output spike-train. As a result, a wider array of spike-time statistics were incorporated in our fits and the accuracy was much higher. Figure 1.1 delineates a schematic for the stochastic neural coder.

Let $s(t) = s$ denote a constant signal input to the neuron and $x[i]$ denote the randomness in the firing threshold of the neuron (see Figure 3.1.) Assume $x[i]$ to be a wide-sense stationary process with a known correlation function $\mathbf{R}_x$.

We can write the reconstructed signal as

$$r(t) = (s + A - \gamma) \exp(-\frac{t}{\tau}) * \sum_i \delta(t - t_i), \tag{3.1}$$

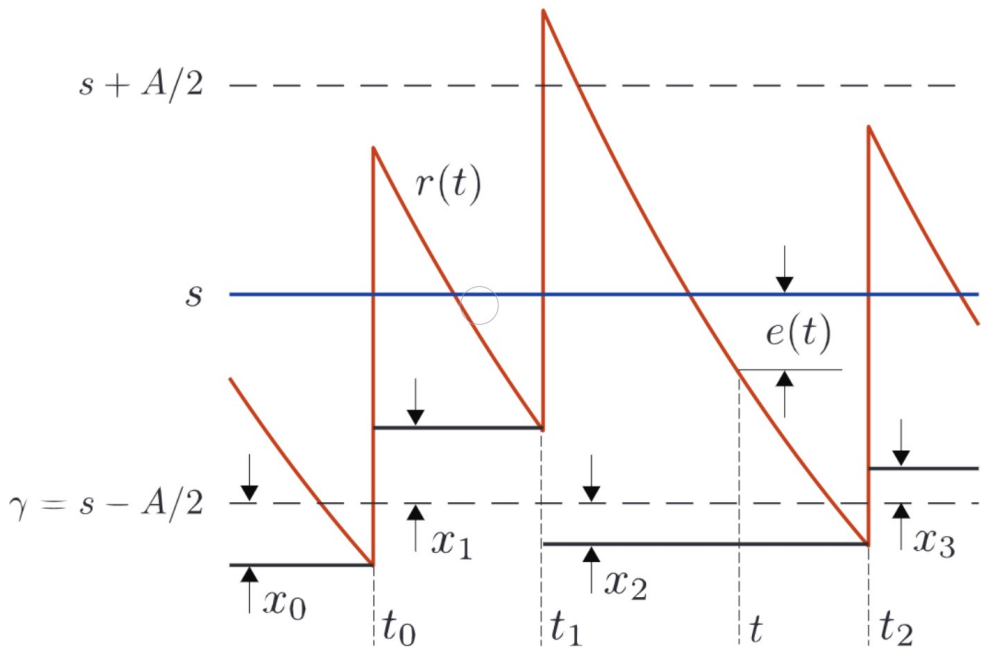

Figure 3.1: The reconstruction signal $r(t)$. At $t = t_0$, when the threshold is crossed, a spike is fired and $r(t)$ jumps by A to a value of $s + A/2 + x_0$. It then linearly decays until the error signal $e(t) = r(t) - s(t)$ crosses the value $A/2 - x_1$.

where γ denotes the instantaneous threshold value, t_i is the i^{th} spike time, τ is the time constant of the reconstruction filter and $x[i]$ represents the threshold perturbation (jitter) at the time of the i^{th} spike. Since x is stationary, the $x[i]$ are identically distributed and are correlated by the function $\mathbf{R}_x$.
Due to wide-sense stationarity of x,

$$\mathbf{R}_x(i, j) = \mathbf{R}_x(j - i). \tag{3.2}$$

Consider $t \in [t_i, t_{i+1}]$:

$$r(t) = (s + A/2 + x[i]) \exp\left(-\frac{t - t_i}{\tau}\right). \tag{3.3}$$

Linearizing the exponential via a first-order Mclaurin series approximation yields:

$$r(t) = (s + A/2 + x[i]) \left(1 - \frac{t - t_i}{\tau}\right). \tag{3.4}$$

The linear function holds from $t = t_i$ to $t = t_{i+1}^-$ (that is, just before the $i + 1^{th}$ spike).

$$r(t_{i+1}^-) = s - (A/2 - x[i+1]) \tag{3.5}$$

$$r(t_{i+1}^-) = (s + A/2 + x[i]) \left(1 - \frac{t_{i+1} - t_i}{\tau}\right) \tag{3.6}$$

$$\Delta_{i+1} = t_{i+1} - t_i = [r(t_{i+1}) - r(t_i)]/[(s + A/2 + x[i])/\tau], \tag{3.7}$$

which assuming linearity, can be written as:

$$\Delta_{i+1} = \frac{[A - x[i+1] + x[i]]}{m}, \tag{3.8}$$

where $m \approx \frac{1}{\tau}\{s + A/2\}$ and $x[i] << s + A/2$.

Using the above results for the ISIs, we shall obtain the SCCs. Since the $x[i]$ all have the same mean,

$$E[\Delta_{i+1}] = \frac{A}{m}, \tag{3.9}$$

$$\mathrm{Cov}[\Delta_{i+1}, \Delta_{i+1}] = E[\{\Delta_{i+1} - E[\Delta_{i+1}]\}^2] \tag{3.10}$$

$$= \frac{E[x[i+1]^2 + x[i]^2 - 2x[i+1]x[i]]}{m^2} \tag{3.11}$$

$$\mathrm{Cov}[\Delta_{i+1}, \Delta_{i+1}] = \frac{2}{m^2}\{\mathbf{R}_x(0) - \mathbf{R}_x(i+1-i)\} \tag{3.12}$$

$$= \frac{2}{m^2}\{\mathbf{R}_x(0) - \mathbf{R}_x(1)\} \tag{3.13}$$

$$\mathrm{Cov}[\Delta_{i+1}, \Delta_{i+2}] = \frac{E[x[i+1]x[i+2] - x[i+2]^2 - x[i]x[i+2] + x[i+1]x[i])]}{m^2} \tag{3.14}$$

$$\mathrm{Cov}[\Delta_{i+1}, \Delta_{i+2}] = \frac{\mathbf{R}_x\left(i+2-(i+1)\right) + \mathbf{R}_x\left(i+1-i\right) - \mathbf{R}_x\left(0\right) - \mathbf{R}_x\left(i+2-i\right)}{m^2} \tag{3.15}$$

$$= \frac{1}{m^2}\left\{2\mathbf{R}_x\left(1\right) - \mathbf{R}_x\left(0\right) - \mathbf{R}_x\left(2\right)\right\}. \tag{3.16}$$

Similarly, we have:

$$\mathrm{Cov}\left[\left(t_{i+1} - t_i\right), \left(t_{i+k+1} - t_{i+k}\right)\right] = \frac{\mathbf{R}_x\left(i+k-i\right) + \mathbf{R}_x\left(i+k+1-(i+1)\right)}{m^2} - \frac{\mathbf{R}_x\left(i+k+1-i\right) + \mathbf{R}_x\left(i+k-(i+1)\right)}{m^2} \tag{3.17}$$

$$= \frac{1}{m^2}\left\{2\mathbf{R}_x\left(k\right) - \mathbf{R}_x\left(k+1\right) - \mathbf{R}_x\left(k-1\right)\right\}. \tag{3.18}$$

Using the above results, we can simplify our expressions for the serial correlation coefficient (ρ_k) of the k^{th} lag as follows:

$$\rho_k = \frac{Cov(\Delta_i, \Delta_{i+k})}{Cov(\Delta_i)}, \tag{3.19}$$

$$= \frac{E[\{\Delta_i - E[\Delta_i]\}\{\Delta_{i+k} - E[\Delta_{i+k}]\}]}{E[(\Delta_i - E[\Delta_i])^2]}. \tag{3.20}$$

$$\rho_0 = 1, \tag{3.21}$$

$$\rho_1 = -\frac{1}{2}\frac{\mathbf{R}_x\left(0\right) - 2\mathbf{R}_x\left(1\right) + \mathbf{R}_x\left(2\right)}{\mathbf{R}_x\left(0\right) - \mathbf{R}_x\left(1\right)}, \tag{3.22}$$

$$\rho_k = -\frac{1}{2}\frac{\mathbf{R}_x(k-1) - 2\mathbf{R}_x(k) + \mathbf{R}_x(k+1)}{\mathbf{R}_x\left(0\right) - \mathbf{R}_x(1)}, \quad k \geq 2. \tag{3.23}$$

3.2 Sum of SCCs

We can now determine $\sum_1^\infty \rho_k$. Let us first transform the ρ_k as follows:

$$z_k = -2\rho_k \left(\mathbf{R}_x(0) - \mathbf{R}_x(1)\right), \quad \text{where } k > 1, \tag{3.24}$$

then z_k is a moving-average process defined as

$$z_k = \mathbf{R}_x(k-1) - 2\mathbf{R}_x(k) + \mathbf{R}_x(k+1), \tag{3.25}$$

from which, it is easy to show that

$$\sum_{k=1}^{N} z_k = \mathbf{R}_x(N+1) - \mathbf{R}_x(N) + \mathbf{R}_x(0) - \mathbf{R}_x(1). \tag{3.26}$$

When $N \to \infty$, $(\mathbf{R}_x(N+1) - \mathbf{R}_x(N)) \to 0$. Thus,

$$\lim_{N\to\infty} \sum_{k=1}^{N} z_k = \mathbf{R}_x(0) - \mathbf{R}_x(1). \tag{3.27}$$

Hence, from Equations 3.24 and 3.27, we obtain the limiting sum of the ISI serial correlation coefficients for the spike-train from a linearized source-coding neuron

$$\sum_{k=1}^{\infty} \rho_k = -\frac{1}{2}. \tag{3.28}$$

Note that the serial correlation coefficients (SCCs) given by Equations 3.21-3.23 are, to the first order, independent of the slope m of the reconstruction filter (the decay rate of the adaptive threshold) and the reconstruction filter gain A. Thus, the observed correlation structure of the spike-train is determined solely by the noise statistics of the signal and the noise statistics of the firing threshold ($A/2+x(t)$). Equation 3.28 explains a very remarkable result observed in the power spectrum density of P-type afferent spike trains, namely a spike train power spectrum with a low frequency block (DC-block). Cox, Lewis and Lewis [46] predicted a DC-block for spike trains having SCC statistics that sum to -0.5. Chacron et al. [47] verified this behavior in their model of spike trains with negative ISI correlations. Thus, the fact that negative ISI correlations decrease the power at low frequencies perfectly describes

the experimentally observed P-type electrosense data and our model of the stochastic neuron coder.

3.3 Spike-time Statistics Assuming Filtered White Noise

In this section, we limit the threshold perturbations to a filtered white noise (Gauss-Markov or Ornstein-Uhlenbeck) process. The stochastic firing threshold could be a result of noisy (probabilistic) transitions between various states in voltage-gated ion channels, resulting in spike-time jitter. Since there are thousands of individual channels present along the axon and at the synapses, the law of large numbers suggests that the aggregate effect on the membrane potential will be approximately Gaussian. Thus, a Gauss-Markov threshold process is a valid assumption. The primary justification for using a Gaussian random process, however, was the observed ISI histograms of experimental data. The ISI histograms are typically unimodally distributed, but certain neurons exhibit bi-modal spreads. A strong attestation to our model is that we were able to re-create these trends by appropriately filtering the Gauss-Markov threshold process without introducing any additional parameters or exceptions.

P-type afferent data display either of two major trends in their ISI and the serial correlation coefficients. The trends we observe in SCCs are:

1. The first lag value is lower than -0.5 and the subsequent lags oscillate before damping down to zero.

2. The first lag value is higher than -0.5 and the subsequent lags are all negative as they approach zero.

We used Equation 3.22 to obtain a mathematical expression for the first lag of the SCCs for a filtered threshold, (note that filtration of the threshold does not yet have a known, explicit biophysical mechanism in the neuron). The two major trends in the SCCs can be reproduced as derived in the subsequent analysis.

3.3.1 Correlated low-pass filtered threshold noise

We begin with a Gaussian noise threshold and proceed to filter it with a first-order RC (resistor-capacitor) low-pass. Denoting the impulse response as $h[n]$ and the filtered threshold correlation as $r[k]$,

$$h[n] = a^n u[n], \tag{3.29}$$

where $|a| < 1$ and $u[n]$ is the discrete Heaviside step function.

$$r[k] = \sum_{n=-\infty}^{\infty} (a^n u[n])(a^{n-k} u[n-k]), \tag{3.30}$$

$$= \sum_{n=k}^{\infty} (a^n)(a^{n-k}), \tag{3.31}$$

$$= a^k \left[\sum_{m=0}^{\infty} (a^2)^m \right], \tag{3.32}$$

$$= a^k \left[\frac{1}{1-a^2} \right]. \tag{3.33}$$

Plugging these results into Equation 3.22, yields

$$\rho_1 = -\frac{1}{2}\{1-a\} \tag{3.34}$$

$$0 \leq a \leq 1 \tag{3.35}$$

$$0 \leq 1-a \leq 1 \tag{3.36}$$

$$\rho_1 \geq -\frac{1}{2}. \tag{3.37}$$

The first lag is greater than -1/2 and $-1/2 < \rho_i < 0$ for $i > 1$, which is one of the two observed forms of SCCs obtained from P-type electric fish data (Figure 5.1).

Additionally, looking at the auto-correlation function for a low-pass filtered Gaussian noise, as shown in Figure 3.2, we see that $\mathbf{R}_x(n) > 0$ for all integer n. Thus, using Equation 3.22:

$$\rho_1 = -\frac{1}{2}\left\{1 + \frac{\mathbf{R}_x(2) - \mathbf{R}_x(1)}{\mathbf{R}_x(0) - \mathbf{R}_x(1)}\right\}, \tag{3.38}$$

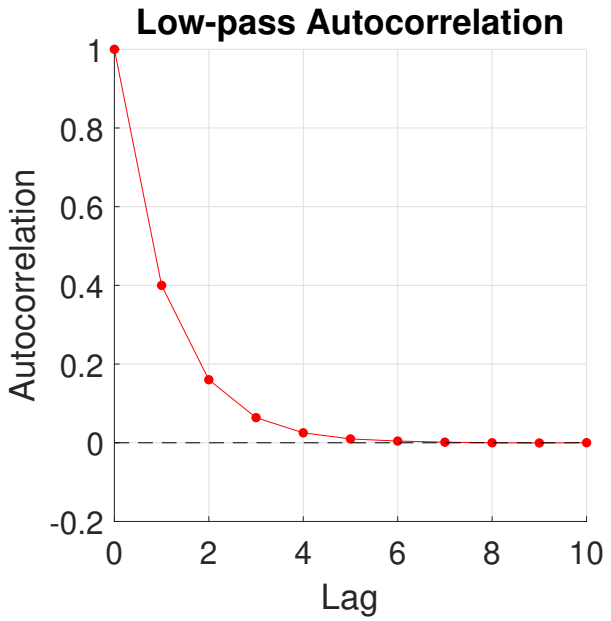

Figure 3.2: The auto-correlation function $\mathbf{R}_x(n)$ where x is a first order, all-pole low-pass filtered Gaussian noise with a = 0.4.

$$\frac{\mathbf{R}_x(2) - \mathbf{R}_x(1)}{\mathbf{R}_x(0) - \mathbf{R}_x(1)} \leq 0. \tag{3.39}$$

Thus, $\rho_1 \geq -\frac{1}{2}$ as expected.

3.3.2 Correlated high-pass filtered noise

Similar to the low-pass case, denote the the first-order high-pass filter response as $h_{hp}[n]$ and an auto-correlation as $r_{hp}[k]$,

$$h_{hp}[n] = (-b)^n u[n], \tag{3.40}$$

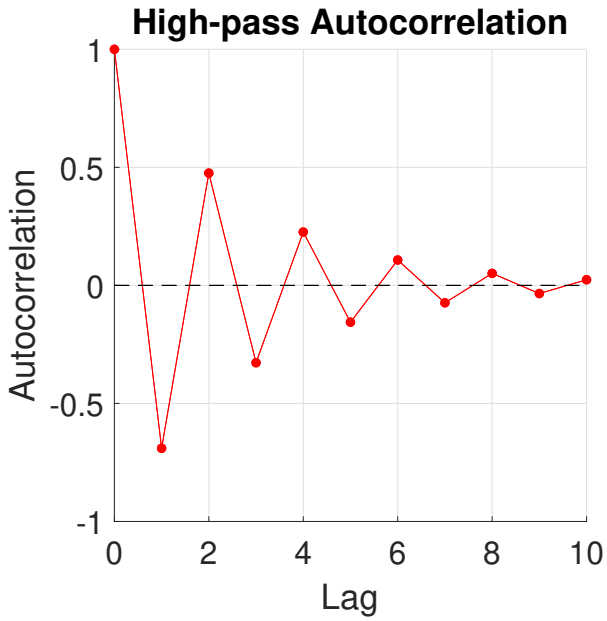

Figure 3.3: The auto-correlation function $\mathbf{R}_x(n)$ where x is a first order, all-pole high-pass filtered Gaussian noise.

where $|b| < 1$, $b > 0$ and $u[n]$ is the discrete Heaviside step function.

$$r_{hp}[k] = \sum_{n=-\infty}^{\infty} ((-b)^n u[n])((-b)^{n-k} u[n-k]), \tag{3.41}$$

$$= \sum_{n=k}^{\infty} ((-b)^n)((-b)^{n-k}), \tag{3.42}$$

$$= (-b)^k \left[\sum_{m=0}^{\infty} (b^2)^m \right], \tag{3.43}$$

$$= (-b)^k \left[\frac{1}{1-b^2} \right]. \tag{3.44}$$

Plugging these results into Equation 3.22, yields

$$\rho_1 = -\frac{1}{2}\{1+b\} \tag{3.45}$$

$$0 \leq b \leq 1 \tag{3.46}$$

$$1 \leq 1+b \leq 2 \tag{3.47}$$

$$\rho_1 \leq -\frac{1}{2}. \tag{3.48}$$

The first lag is lesser than -1/2 and subsequent lags alternate between positive and negative as we observe in experimental data (Figure 5.3).

Additionally, we see in Figure 3.3 that the $\mathbf{R}_x(n)$ exhibit alternation between positive and negative values as theoretically predicted. Considering ρ_1,

$$\rho_1 = -\frac{1}{2}\left\{1 + \frac{\mathbf{R}_x(2) - \mathbf{R}_x(1)}{\mathbf{R}_x(0) - \mathbf{R}_x(1)}\right\}, \tag{3.49}$$

because $\mathbf{R}_x(0)$, $\mathbf{R}_x(2)$ are positive while $\mathbf{R}_x(1)$ is negative,

$$\frac{\mathbf{R}_x(2) - \mathbf{R}_x(1)}{\mathbf{R}_x(0) - \mathbf{R}_x(1)} \geq 0, \tag{3.50}$$

which also yields $\rho_1 \leq -\frac{1}{2}$.

When the bandwidth and cut-off frequency of the filter are adjusted appropriately, we are able to achieve a bi-modality in the ISI histograms of the simulated spike trains, even though the original input noise to the filter was a uni-modal Gaussian noise. This is also observed in the P-type afferent data and hence strongly reinforces the theoretical idea of having a filtered noise as the threshold perturbation.

CHAPTER 4

GROWTH IN VARIANCE

4.1 Limitations of Linear Approximation

In Chapter 3, our theoretical derivations were based on a linear approximation of the exponentially decaying reconstruction. Although successful in explaining some major trends in data, a closer look at some finer details pointed out some significant differences from experimental data. Most importantly, our previous finding that the sum of SCCs should be exactly -0.5 is not necessarily the case. We show here that the memory leakage in reconstruction and the growth in variance are strongly linked to the sum of SCCs deviating from the idealistic -0.5. In the current section, we re-derive our expressions with a tighter approximation which showcases the "imperfections" in the sum of SCCs and yields the growth of variance that we observe in the experimental data.

4.2 Logarithmic Approximation Using Small Noise

Let $s(t) = s$ denote a constant signal input to the neuron and $x[n]$ denote the randomness in the firing threshold of the neuron. Assume $x[n]$ to be a zero mean, wide-sense stationary process with a known correlation function $\mathbf{R}_x[n]$:

$$\mathbf{R}_x[m, n+m] = \mathbf{R}_x[n]. \tag{4.1}$$

We can write the reconstructed signal as:

$$r(t) = (s + A - \gamma) \exp\left(-\frac{t}{\tau}\right) * \sum_i \delta(t - t_i), \tag{4.2}$$

where γ is the instantaneous threshold, and t_i is the i^{th} spike time; $x[i]$ represents the threshold perturbation (jitter) at the time of the i^{th} spike. Since $x[n]$ is stationary, the $x[i]$ are identically distributed and are correlated according to the function $\mathbf{R}_x$.

Consider $t \in [t_i, t_{i+1})$, :

$$r(t) = (s + A/2 + x[i]) \exp\left(-\frac{t - t_i}{\tau}\right) \tag{4.3}$$

at $t = t_{i+1}$

$$r(t_{i+1}^-) = (s + A/2 + x[i]) \exp\left(-\frac{t_{i+1} - t_i}{\tau}\right). \tag{4.4}$$

Also, we know that spike time t_{i+1} occurs at the instant that the error signal exactly equals the threshold $A/2 + x[i+1]$.

$$r(t_{i+1}^-) = s - A/2 + x[i+1]. \tag{4.5}$$

Using Equations 4.4 and 4.5 and rearranging, gives us an expression for the inter-spike interval (ISI),

$$t_{i+1} - t_i = \tau \left\{ ln\left(\frac{s + A/2 + x[i]}{s - A/2 + x[i+1]}\right)\right\}. \tag{4.6}$$

Letting $\alpha = 1/(s + A/2)$, $\beta = 1/(s - A/2)$ and approximating $ln(1 + x) = x$ gives:

$$t_{i+1} - t_i = \tau \left\{ ln\left(\frac{\beta(1 + \alpha x[i])}{\alpha(1 + \beta x[i+1])}\right)\right\} \tag{4.7}$$

$$\Delta_{i+1} = \tau \left\{ ln\left(\frac{\beta}{\alpha}\right) - \beta x[i+1] + \alpha x[i]\right\} \tag{4.8}$$

$$E[\Delta_{i+1}] = \tau \left\{ ln\left(\frac{\beta}{\alpha}\right)\right\} \tag{4.9}$$

$$\mathrm{Cov}(\Delta_k, \Delta_{k+l}) = E[\Delta_k - E[\Delta_k]][\Delta_{k+l} - E[\Delta_{k+l}]] \tag{4.10}$$

$$\mathrm{Cov}(\Delta_k, \Delta_{k+l}) = (\alpha^2 + \beta^2)\mathbf{R}_x(l) - \alpha\beta(\mathbf{R}_x(l+1) + \mathbf{R}_x(l-1)). \tag{4.11}$$

Also note that

$$\mathrm{Cov}(\Delta_k, \Delta_k) = \sigma_1^2 \tag{4.12}$$

$$\sigma_1^2 = (\alpha^2 + \beta^2)\mathbf{R}_x(0) - 2\alpha\beta(\mathbf{R}_x(1)). \tag{4.13}$$

This gives us the serial correlation coefficients at the k^{th} lag as

$$\rho_0 = 1, \tag{4.14}$$

$$\rho_1 = \frac{(\alpha^2 + \beta^2)\mathbf{R}_x(1) - \alpha\beta(\mathbf{R}_x(0) + \mathbf{R}_x(2))}{(\alpha^2 + \beta^2)\mathbf{R}_x(0) - 2\alpha\beta(\mathbf{R}_x(1))}, \tag{4.15}$$

$$\rho_k = \frac{(\alpha^2 + \beta^2)\mathbf{R}_x(k) - \alpha\beta(\mathbf{R}_x(k-1) + \mathbf{R}_x(k+1))}{(\alpha^2 + \beta^2)\mathbf{R}_x(0) - 2\alpha\beta(\mathbf{R}_x(1))}, \quad k \geq 2. \tag{4.16}$$

4.3 Growth in Variance of ISIs

We shall now take a look at higher order intervals and how the variance grows with respect to the order of the interval

$$\Delta_i^k = t_{i+k+1} - t_{i+1}. \tag{4.17}$$

Denote the k^{th} order variance as:

$$\sigma_k^2 = E[(\Delta_i^k - E[\Delta_i^k])^2]. \tag{4.18}$$

Expanding Equation 4.18 using Equations 4.17 and 3.20, we can show that

$$\sigma_k^2 = k\sigma_1^2 \left\{ 1 + 2\sum_{l=1}^{k-1}(1 - \frac{l}{k})\rho_l \right\} \tag{4.19}$$

$$= k\sigma_1^2 \left\{ 1 + 2\sum_{l=1}^{k-1}\rho_l \right\} - \sigma_1^2 \left\{ 2\sum_{l=1}^{k-1} l\rho_l \right\}. \tag{4.20}$$

Hence, we can study the growth in variance of intervals for a spike train by using the SCCs in the above expression, which has constant and linear terms in k. We can see from Equation 4.20 that when the sum of SCCs is precisely -0.5, the linear term in k disappears and we see no growth in variance with respect to k.

4.4 Sum of SCCs

The sum of SCCs has interesting implications, and a closed-form expression provides us with a few insights. Let ρ_l denote the l^{th} serial correlation coefficient (SCC),

$$\sum_{l=1}^{\infty} \rho_l = \sum_{l=1}^{\infty} \frac{1}{2} \frac{(\alpha^2+\beta^2)\mathbf{R}_x(l) - \alpha\beta(\mathbf{R}_x(l-1) + \mathbf{R}_x(l+1))}{((\alpha^2+\beta^2)\mathbf{R}_x(0) - 2\alpha\beta(\mathbf{R}_x(1)))}. \tag{4.21}$$

Simplifying further we obtain:

$$\sum_{l=1}^{\infty} \rho_l = -\frac{1}{2} + \frac{1}{2}(\alpha-\beta)^2 \left[\frac{\mathbf{R}_x(0) + 2S^*}{(\alpha^2+\beta^2)\mathbf{R}_x(0) - 2\alpha\beta\mathbf{R}_x(1)} \right], \tag{4.22}$$

where $\alpha = 1/(s + A/2)$, $\beta = 1/(s - A/2)$ and $S^* = \sum_{n=1}^{\infty} \mathbf{R}_x(n)$. It has been shown that a sum of SCCs = -1/2 corresponds to a perfect DC block [47], which is what we obtain for a white noise, high spike rate assumption for the model. This is a simple consequence of the relation between the power spectral density S_{xx} and the correlation function $\mathbf{R}_x$:

$$\mathcal{F}[\mathbf{R}_x] = S_{xx} \tag{4.23}$$

$$S_{xx}(0) = \sum_{n=-\infty}^{\infty} \mathbf{R}_x(n), \tag{4.24}$$

$$= 2\sum_{n=1}^{\infty} \mathbf{R}_x(n) + \mathbf{R}_x(0), \tag{4.25}$$

$$= 2(-0.5) + 1, \tag{4.26}$$

$$= 0. \tag{4.27}$$

Also note that the LHS in Equation 4.26 is identical to the second term in Equation 4.22. Thus, for a perfect DC block system we can conversely expect the sum of SCCs to be precisely -0.5.

We attribute the deviation of the sum of SCCs from -0.5 to the memory leakage of the stochastic source coding model. Not only does this allow us to explain the growth in variance (previously assumed flat in Chapter 3), but

it also provides a potential method for deriving our model parameters, such as the reconstruction filter's time constant τ and the filter characteristics of the threshold noise, from experimental data.

CHAPTER 5

COMPARISON TO EXPERIMENTAL DATA

The stochastic source-coding neuron was compared to the experimentally obtained spike trains from a brown ghost knifefish (Apteronotus leptorhynchus), which is a species of weakly electric fish in the family Apteronotidae. Afferent nerve fibers carry stimuli to a particular region of the brain. P-type afferents fire with a per-cycle firing probability [48] that depends on the strength of the external stimulus, making them ideal for investigating neural coding when the stimulus is known. The stochastic source-coding neuron was used to predict spike times from P-type afferents with long-term baseline activity. Statistics and reconstructed waveforms were computed for each spike train. We found that the ISI distributions, interval correlations, and spike-time reliability of experimental data were accurately predicted by our stochastic source-coding neuron.

5.1 Spike-Train Statistics

ISI histogram, joint-ISI histogram, SCCs and the variance of the k^{th} interval are the four major statistics we considered for comparing our model to experimental data. We observed that the available data was either uni-modal or bi-modal. The bimodality was often accompanied by very oscillatory SCCs. We hypothesized and subsequently observed that we could obtain both trends by using different filters on a Gaussian noise threshold. Once we "matched" the ISI distribution and the SCC statistics, the joint-ISIs tended to follow suit. The low-pass filtered threshold yielded a uni-modal ISI distribution and a centered joint distribution, plotted in Figure 5.1. The bi-modal ISIs were obtained through a high-pass digital filter. The joint distribution for the high-pass case also differed from the low-pass or uni-modal data as shown in Figure 5.2.

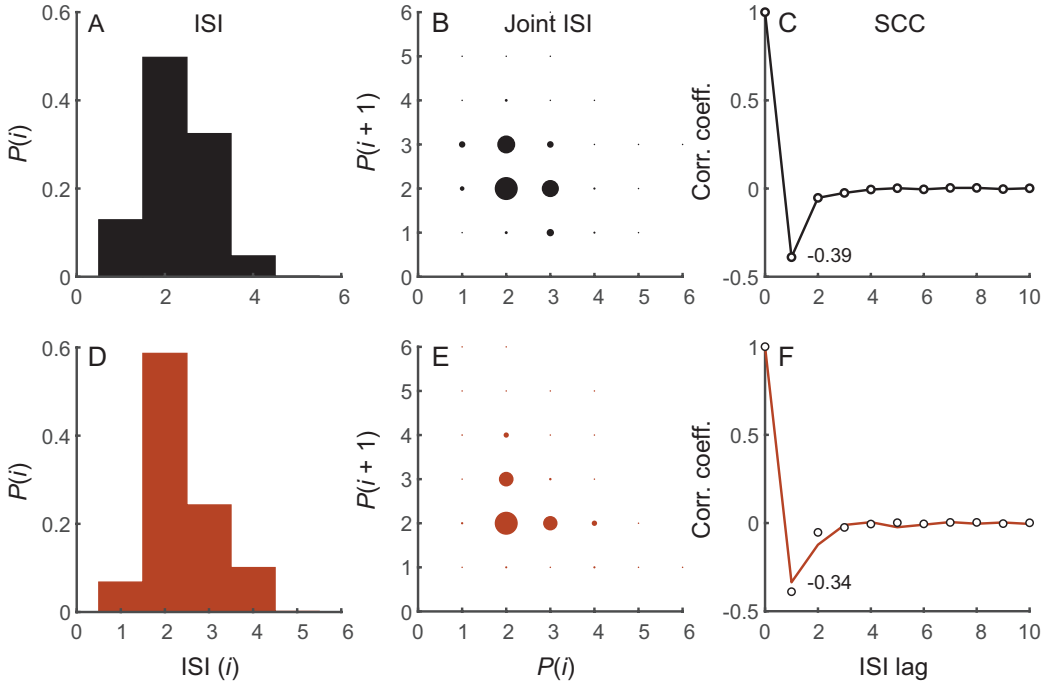

Figure 5.1: Modelling of experimental spike-time data using a low-pass filtered noisy threshold. A shows the inter-spike interval histogram for the experimental data, B is the joint ISI histogram, and C shows the serial correlation coefficients. The plots in red are the corresponding statistics for our stochastic source-coding neuron (SSCN) model. D is the ISI histogram for the SSCN, E represents the joint ISI histogram, and F shows the simulated SCCs overlapped on the experimental SCCs. The first column displays the ISI histogram binned at EOD cycles. The electric fish data has an EOD of around 750 Hz. We used a reconstruction τ of 30 ms for this fit. The pole location for the first-order low-pass filter was 0.4.

The ISIs in the experimental data are correlated (Figure 5.1) and demonstrate long-short anti-correlations (see [3]). The source-coding neuron clearly depicts this long-short anti-correlation pattern. The serial correlation coefficients for the ISI sequence are another way of showing the dependencies in the intervals (Figure 5.1). The P-type and our stochastic source-coding neuron both show strong negative correlations at the first lag (ρ_1 around -0.5) . The source-coding neuron was also able to, almost perfectly, match experimentally observed SCCs as shown in Figure 5.1.

The statistics of the P-type afferents varied considerably over the 53 units (available experimental data) with baseline activity which were studied. On average, the power of the added noisy threshold additive noise power, which best matched the P-type afferent spike train statistics, was relatively small

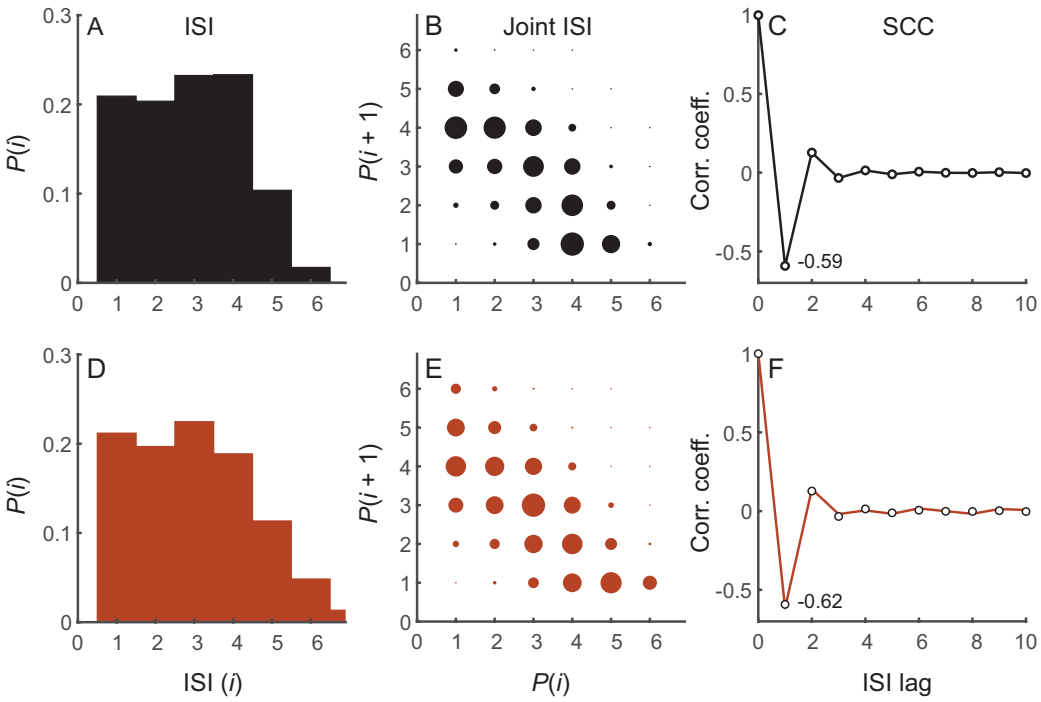

Figure 5.2: Modelling of experimental spike-time data using a high-pass filtered noisy threshold. This particular data sample represents the damped-oscillatory trend of SCCs observed in electric fish data. A shows the inter-spike interval histogram for the experimental data, B is the joint ISI histogram, and C shows the serial correlation coefficients. The plots in red are the corresponding statistics for our stochastic source-coding neuron (SSCN) model. D is the ISI histogram for the SSCN, E represents the joint ISI histogram, and F shows the simulated SCCs overlapped on the experimental SCCs. The displayed ISI and joint ISI histograms are binned at the EOD frequency (750 Hz). We used a reconstruction τ of 26 ms for this fit. The pole location for the first-order high-pass filter was -0.29.

compared to the baseline signal level (-17dB). All observed units showed a negative correlation between succeeding ISIs (that is $\rho_1 < 0$) and none of them showed a positive correlation at the first lag. The experimental data can be primarily divided into two types based on the SCCs.

1. $-0.5 < \rho_1 < 0$, followed by less negative coefficients ρ_k, $k \geq 2$ at higher lags, gradually increasing to zero (Figure 5.1).

2. $-1 < \rho_1 < -0.5$, followed by alternating positive and negative ρ_k, $k \geq 2$, gradually diminishing to zero (Figure 5.3).

The source-coding neuron is able to accurately recreate these two major trends observed in the experimental data. Furthermore, we were able to

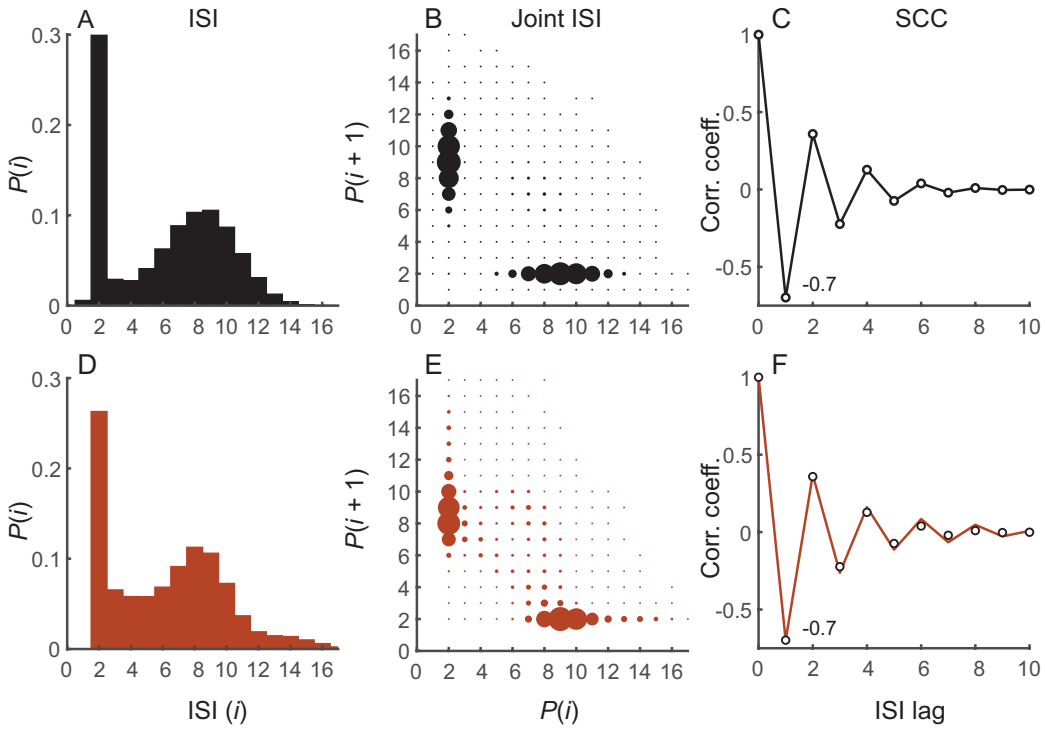

Figure 5.3: Reconstruction of experimental spike-time data using a high-pass filtered noisy threshold. This particular data sample represents a highly-oscillatory trend of SCCs observed in electric fish data. A shows the inter-spike interval histogram for the experimental data, B is the joint ISI histogram, and C shows the serial correlation coefficients. The plots in red are the corresponding statistics for our stochastic source-coding neuron (SSCN). D is the ISI histogram for the SSCN, E represents the joint ISI histogram, and F shows the simulated SCCs overlapped on the experimental SCCs. The displayed ISI and joint ISI histograms are binned at the EOD frequency (750 Hz). We used a reconstruction τ of 60 ms for this fit. The pole location for the first-order high-pass filter was -0.69. Compared to the damped SCC sample, the fit for the highly oscillatory samples tend to have a narrower (higher Q) filter response.

find particular model parameters that yielded statistics almost identical to experimental data as shown in Figures 5.1, 5.2 and 5.3.

The source-coding neuron is also able to replicate the power spectral density observed in experimental spike-trains. These spike trains exhibit a high-pass characteristic, as shown by Johnson [12] in Figure 5.4, in contrast to the renewal process and linear integrate and fire (LIAF) neuron model, the commonly considered standards which have much flatter spectra. This bolsters

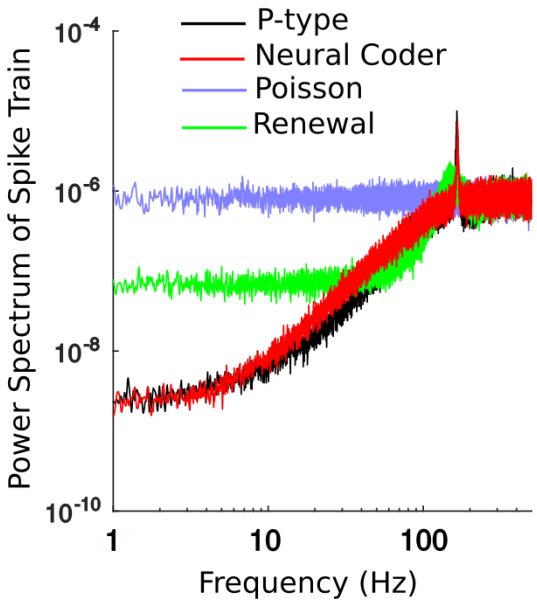

Figure 5.4: Power spectral density of spike trains for various models and experimental data. Note the power spectrum at low frequencies for experimental P-type and the neural source-coding model.

support for P-type afferents exhibiting non-renewal statistics. It has been suggested that power spectra with a high-pass characteristic are related to the signal-processing concept of noise shaping [49]. Noise shaping can improve the resulting signal-to-noise ratio of a quantizer. The high-pass spectra are also closely related to the anti-correlated intervals observed in Figure 5.3.

5.2 Model Parameters

Our stochastic source-coding neuron model primarily uses only four parameters to fit the experimental data. We used the reconstruction filter time constant τ, the base spike rate R and a specific filter for the threshold noise to achieve the fits. The filters chosen were first order low-pass and first order high-pass digital filters and so the filter design needed only two variables, the filter gain and the pole of the filter. It was observed that the SCC behavior is mostly influenced by the pole location while the ISI and joint ISI anti-correlations are largely determined by the reconstruction τ (although there is some coupling between all the parameters).

1. For the first type of SCC behavior, we used a low-pass filtered noise as the threshold with a pole at 0.4 and a reconstruction τ value of 30 ms (Figure 3.2).

2. For the damped oscillatory SCC type of neurons, we used a high-pass filter with a pole at -0.29 and a reconstruction τ of 26 ms (Figure 5.2).

3. For the highly oscillatory form, we used a high-pass filter with a pole at -0.69 and a reconstruction τ of 60 ms (Figure 5.3).

Another metric to look at is the sum of SCCs. It is a good indicator of how rapidly the variance of the k^{th} ISI grows. The closer the value is to -0.5, the flatter the curve. Preliminary investigation has shown that we can directly obtain the reconstruction τ from the SCCs (or sum of SCCs) of an experimental spike train. This is significant because it allows us to artificially produce spike-trains of desired length from data that would otherwise be statistically insufficient for some experiments.

5.3 Growth in Variance

Using the better approximation from Chapter 4, our stochastic source-coding model gives a variance growth in contrast to the linear approximation that yields no growth at all. (Since the sum of SCCs is exactly -0.5, the asymptotic slope from equation is zero (refer to Section 4.4) which does not agree with experimental results.) Using the parameters we used in the previous sections to get a close fit for experimental data, we were also able to obtain accurate slopes for growth in variance for the same neurons. The low-pass noisy threshold neuron shown in Figure 5.1 has a growth of variance plotted in Figure 5.5) and the high-pass neuron (Figure 5.3) model's variance growth is shown in Figure 5.6

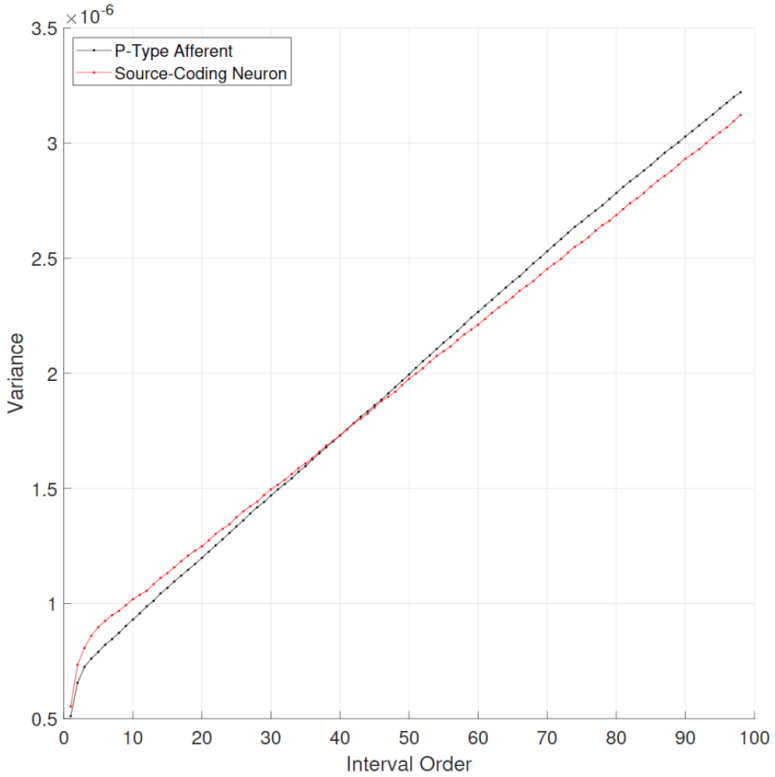

Figure 5.5: Plot of the growth in variance. This particular data sample corresponds to the low-pass fit used in Figure 5.1. The model parameters were exactly the same as those used to generate Figure 5.1.

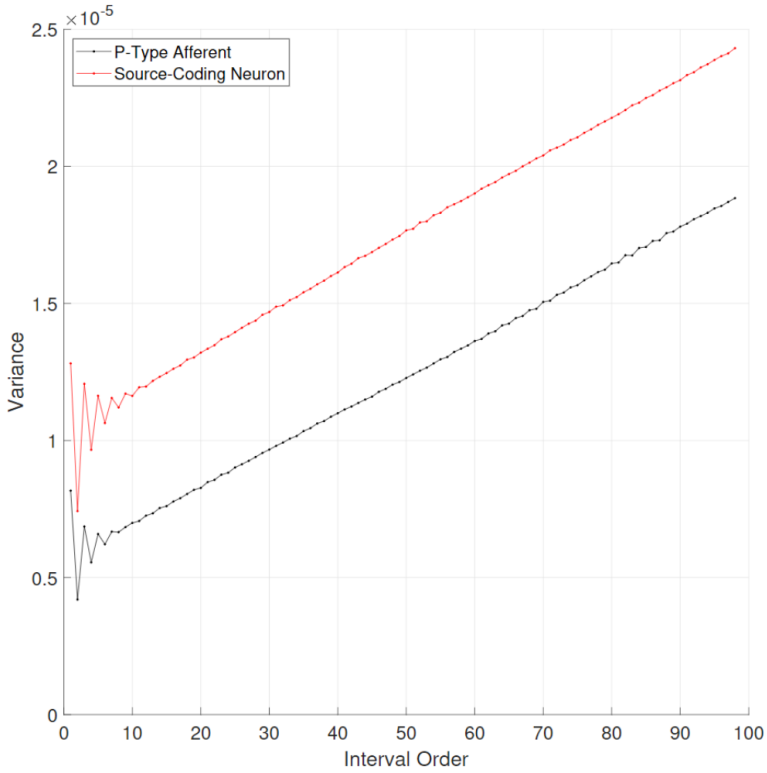

Figure 5.6: Plot of the growth in variance. This particular data sample corresponds to the high-pass fit used in Figure 5.3. The model parameters were exactly the same as those used to generate Figure 5.3.

CHAPTER 6

CONCLUSION

This work provides a stochastic extension to the source-coding neuron [1] by introducing a noisy threshold to jitter the output spike times. The deterministic neural source-coder optimizes metabolic expenditure of spiking and high fidelity for a given stimulus. To account for variability in observed spike-trains, a stochastic component was added to the error threshold responsible for spike times. The signal-to-noise ratio for the threshold was quite high (17 dB) allowing us to make small-signal approximation in the analysis without sacrificing much precision. Even such a small noise power was able to create significant deviation in the spike-time statistics in contrast to those obtained from a deterministic source-coder.

We presume that this stochastic firing threshold is a result of noisy (probabilistic) transitions between various states in voltage-gated ion channels, resulting in spike-time jitter [50]. Due to the thousands of individual channels present in the spike initiation zone [51], the law of large numbers suggests that the aggregate effect on the membrane potential will be approximately Gaussian. Thus, we assumed that the threshold noise was a low-pass or high-pass filtered Gauss-Markov process.

The total deviation in spiking threshold from the asymptotic fixed-threshold $A/2$ is thus $x[n]$. These deviations are small compared to the membrane potential, similar to deviations shown by Fontaine et al. [52]. The presence of weak noise is also compatible with Schwalger and Lindner [53], who reported that weak noise is sufficient to reproduce similar patterns of ISI correlations. It should be noted that the observed ISI anti-correlations are a property of the source-coding mechanism (i.e., the adaptive threshold) and are not due to a noisy threshold. This is because the source-coding neuron encodes the error, i.e., difference between a signal and its estimate (reconstruction). The various patterns of ISI correlations are governed by the characteristics of the noisy threshold. While a low-pass filtered additive noise source (an

Ornstein-Uhlenbeck process) could reproduce observed anti-correlations that were restricted to the range $0 > \rho > -.5$, ISI correlations where $\rho < -.5$ could be reproduced only with a high-pass noise source centered at a frequency precisely half that of the average neuron spike rate. These units tend to have a bi-modal ISI histogram. The high-pass process introduces a very pronounced short-long ISI pattern (Figure 5.3).

In addition to accurately explaining anti-correlations, our model is also able to match the growth in variance as observed from experimental data. This can be shown to have implications for weak signal detection [3] and disproves the belief that such spike-trains follow renewal statistics (Poisson spike-trains for instance have a much higher slope).

The sum of SCCs is also an interesting metric that allows us to quantify the memory loss of the source-coding neuron. A linear approximation of the reconstruction yields the sum of SCCs to be exactly -0.5 and thus a zero asymptotic variance slope. As we see the sum deviate from -0.5 (the sum of SCCs is always greater than -0.5 as shown in Equation 4.22) the variance growth is that much more rapid.

Future work can involve showcasing a better weak-signal detection using our SSCN model spike-train compared to a Poisson spike-train. It would also be worthwhile to be able to generate the model parameters directly from experimental data such as the reconstruction filter time constant τ and the threshold noise profile. Recreating spike-trains in such fashion opens up a whole new world of data that is statistically identical to experimental spike-trains and will prove extremely helpful for a wide variety of studies.

REFERENCES

[1] E. C. Johnson, D. L. Jones, and R. Ratnam, "A minimum-error, energy-constrained neural code is an instantaneous-rate code," *Journal of Computational Neuroscience*, vol. 40, no. 2, pp. 193–206, 2016.

[2] G. L. Gerstein and B. Mandelbrot, "Random walk models for the spike activity of a single neuron," *Biophysical Journal*, vol. 4, no. 1 Pt 1, p. 41, 1964.

[3] R. Ratnam and M. E. Nelson, "Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals," *The Journal of Neuroscience*, vol. 20, no. 17, pp. 6672–6683, 2000.

[4] S. Hagiwara, "Analysis of interval fluctuation of the sensory nerve impulse." *The Japanese Journal of Physiology*, vol. 4, no. 3, pp. 234–240, 1954.

[5] M. J. Chacron, A. Longtin, and L. Maler, "Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli," *The Journal of Neuroscience*, vol. 21, no. 14, pp. 5328–5343, 2001.

[6] J. B. Goense, R. Ratnam, and M. E. Nelson, "Burst firing improves the detection of weak signals in spike trains," *Neurocomputing*, vol. 52, pp. 103–108, 2003.

[7] J. B. Goense and R. Ratnam, "Continuous detection of weak sensory signals in afferent spike trains: The role of anti-correlated interspike intervals in detection performance," *Journal of Comparative Physiology A*, vol. 189, no. 10, pp. 741–759, 2003.

[8] F. Farkhooi, M. F. Strube-Bloss, and M. P. Nawrot, "Serial correlation in neural spike trains: Experimental evidence, stochastic modeling, and single neuron variability," *Physical Review E*, vol. 79, no. 2, p. 021905, 2009.

[9] R. Brandman and M. E. Nelson, "A simple model of long-term spike train regularization," *Neural Computation*, vol. 14, no. 7, pp. 1575–1597, 2002.

[10] M. J. Chacron, K. Pakdaman, and A. Longtin, "Interspike interval correlations, memory, adaptation, and refractoriness in a leaky integrate-and-fire model with threshold fatigue," *Neural Computation*, vol. 15, no. 2, pp. 253–278, 2003.

[11] D. L. Jones, E. C. Johnson, and R. Ratnam, "A stimulus-dependent spike threshold is an optimal neural coder," *Frontiers in Computational Neuroscience*, vol. 9, p. 61, 2015.

[12] E. C. Johnson, "Minimum-error, energy-constrained source coding by sensory neurons," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2016.

[13] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain," *Journal of Cerebral Blood Flow & Metabolism*, vol. 21, no. 10, pp. 1133–1145, 2001.

[14] A. Hasenstaub, S. Otte, E. Callaway, and T. J. Sejnowski, "Metabolic cost as a unifying principle governing neuronal biophysics," *Proceedings of the National Academy of Sciences*, vol. 107, no. 27, pp. 12 329–12 334, 2010.

[15] S. B. Laughlin, R. R. de Ruyter van Steveninck, and J. C. Anderson, "The metabolic cost of neural information," *Nature Neuroscience*, vol. 1, no. 1, pp. 36–41, 1998.

[16] S. B. Laughlin, "Energy as a constraint on the coding and processing of sensory information," *Current Opinion in Neurobiology*, vol. 11, no. 4, pp. 475–480, 2001.

[17] P. Lennie, "The cost of cortical computation," *Current Biology*, vol. 13, no. 6, pp. 493–497, 2003.

[18] J. E. Niven and S. B. Laughlin, "Energy limitation as a selective pressure on the evolution of sensory systems," *Journal of Experimental Biology*, vol. 211, no. 11, pp. 1792–1804, 2008.

[19] B. Sengupta, M. Stemmler, S. B. Laughlin, and J. E. Niven, "Action potential energy efficiency varies among neuron types in vertebrates and invertebrates," *PLoS Computational Biology*, vol. 6, no. 7, p. e1000840, 2010.

[20] E. C. Johnson, D. L. Jones, and R. Ratnam, "Minimum squared-error, energy-constrained encoding by adaptive threshold models of neurons," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2015, pp. 1337–1341.

[21] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[22] R. Baddeley, L. F. Abbott, M. C. Booth, F. Sengpiel, T. Freeman, E. A. Wakeman, and E. T. Rolls, "Responses of neurons in primary and inferior temporal visual cortices to natural scenes," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 264, no. 1389, pp. 1775–1783, 1997.

[23] W. B. Levy and R. A. Baxter, "Energy efficient neural codes," *Neural Computation*, vol. 8, no. 3, pp. 531–543, 1996.

[24] P. Crotty and W. B. Levy, "Energy-efficient interspike interval codes," *Neurocomputing*, vol. 65, pp. 371–378, 2005.

[25] T. Berger and W. B. Levy, "A mathematical theory of energy efficient neural computation and communication," *IEEE Transactions on Information Theory*, vol. 56, no. 2, pp. 852–874, 2010.

[26] J. Xing, T. Berger, and T. J. Sejnowski, "A Berger-Levy energy efficient neuron model with unequal synaptic weights," *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 2964–2968, 2012.

[27] J. Xing and T. Cov, "Energy efficient neurons with generalized inverse gaussian conditional and marginal hitting times," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2013, pp. 1824–1828.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[29] Y.-H. Liu and X.-J. Wang, "Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron," *Journal of Computational Neuroscience*, vol. 10, no. 1, pp. 25–45, 2001.

[30] R. Kobayashi, Y. Tsubo, and S. Shinomoto, "Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold," *Frontiers in Computational Neuroscience*, vol. 3, 2009.

[31] J. Shin, C. Koch, and R. J. Douglas, "Adaptive neural coding dependent on the time-varying statistics of the somatic input current," *Neural Computation*, vol. 11, no. 8, pp. 1893–1913, 1999.

[32] J. Shin, "Adaptation in spiking neurons based on the noise shaping neural coding hypothesis," *Neural Networks*, vol. 14, no. 6, pp. 907–919, 2001.

[33] J. Shin, "A unifying theory on the relationship between spike trains, EEG, and ERP based on the noise shaping/predictive neural coding hypothesis," *Biosystems*, vol. 67, no. 1, pp. 245–257, 2002.

[34] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*. IEEE Press, Piscataway, NJ, 2005.

[35] M. N. Shadlen and W. T. Newsome, "The variable discharge of cortical neurons: implications for connectivity, computation, and information coding," *The Journal of Neuroscience*, vol. 18, no. 10, pp. 3870–3896, 1998.

[36] R. R. de Ruyter van Steveninck, G. D. Lewen, S. P. Strong, R. Koberle, and W. Bialek, "Reproducibility and variability in neural spike trains," *Science*, vol. 275, no. 5307, pp. 1805–1808, 1997.

[37] J. Von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," *Automata Studies*, vol. 34, pp. 43–98, 1956.

[38] L. Schuchman, "Dither signals and their effect on quantization noise," *IEEE Transactions on Communication Technology*, vol. 12, no. 4, pp. 162–165, December 1964.

[39] K. Wiesenfeld, F. Moss et al., "Stochastic resonance and the benefits of noise: from ice ages to crayfish and squids," *Nature*, vol. 373, no. 6509, pp. 33–36, 1995.

[40] M. D. McDonnell and D. Abbott, "What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology," *PLoS Computational Biology*, vol. 5, no. 5, p. e1000348, 2009.

[41] H. Chen, L. R. Varshney, and P. K. Varshney, "Noise-enhanced information systems," *Proceedings of the IEEE*, vol. 102, no. 10, pp. 1607–1621, 2014.

[42] A. Longtin, "Stochastic resonance in neuron models," *Journal of Statistical Physics*, vol. 70, no. 1-2, pp. 309–327, 1993.

[43] J. Collins, C. C. Chow, T. T. Imhoff et al., "Stochastic resonance without tuning," *Nature*, vol. 376, no. 6537, pp. 236–238, 1995.

[44] D. R. Chialvo, A. Longtin, and J. Müller-Gerking, "Stochastic resonance in models of neuronal ensembles," *Physical Review E*, vol. 55, no. 2, p. 1798, 1997.

[45] J. E. Levin and J. P. Miller, "Broadband neural encoding in the cricket cereal sensory system enhanced by stochastic resonance," *Nature*, vol. 380, no. 6570, pp. 165–168, 1996.

[46] D. Cox, P. A. Lewis, and P. A.W. Lewis, *The Statistical Analysis of Series of Events*, Hethuen, London, 1966.

[47] M. J. Chacron, B. Lindner, and A. Longtin, "ISI correlations and information transfer," *Fluctuation and Noise Letters*, vol. 4, no. 01, pp. L195–L205, 2004.

[48] M. E. Nelson, Z. Xu, and J. Payne, "Characterization and modeling of p-type electrosensory afferent responses to amplitude modulations in a wave-type electric fish," *Journal of Comparative Physiology A*, vol. 181, no. 5, pp. 532–544, 1997.

[49] O. Ávila Åkerberg and M. J. Chacron, "Nonrenewal spike train statistics: causes and functional consequences on neural coding," *Experimental Brain Research*, vol. 210, no. 3-4, pp. 353–371, 2011.

[50] J. A. White, J. T. Rubinstein, and A. R. Kay, "Channel noise in neurons," *Trends in Neurosciences*, vol. 23, no. 3, pp. 131–137, 2000.

[51] B. Hille, *Ion Channels of Excitable Membranes*. Sinauer Sunderland, MA, 2001.

[52] B. Fontaine, J. L. Peña, and R. Brette, "Spike-threshold adaptation predicted by membrane potential dynamics in vivo," *PLoS Computational Biology*, vol. 10, no. 4, p. e1003560, 2014.

[53] T. Schwalger and B. Lindner, "Patterns of interval correlations in neural oscillators with adaptation," *Frontiers in Computational Neuroscience*, vol. 7, no. 164, 2013.