Realistic spiking neuron statistics in a population are described by a single parametric distribution

Lauren E. Crow[†]
Sponsor: Cheng Ly[†]

[†]*Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, Richmond, VA 23284-3083 U.S.A.*
∗ E-mail: crowle@vcu.edu ; CLy@vcu.edu

Abstract

The spiking of activity of neurons throughout the cortex is random and complicated. This complicated activity requires theoretical formulations in order to understand the underlying principles of neural processing. A key aspect of theoretical investigations is characterizing the probability distribution of spiking activity. This study aims to better understand the statistics of the time between spikes, or interspike interval, in both real data and a spiking model with many time scales. Exploration of the interspike intervals of neural network activity can provide a better understanding of neural responses to different stimuli. We consider different parametric distribution fitting techniques to characterize the random spike times of a population of neurons in the visual cortex of a mammal. Five different probability distribution functions were considered, including three mixture models, and their goodness of fit was determined through two criteria: maximum likelihood and Akaike Information Criteria. Despite being largely heterogeneous, both criteria indicated that one distribution, although different for each criteria, was the best fitting for all of the neurons in the data set. The Gamma-Gamma mixture distribution was the best according to maximum likelihood and the Exponential distribution was the best according to AIC. The statistical methodology applied to a burst model yielded the same results, and the AIC formula was further investigated to better understand its consistent selection of the same parametric distribution. We find that complicated neural spiking activity can sometimes be described by a single parametric distribution, which is hopefully comforting for theorists.

1 Introduction

Cortical neural activity in the form of spikes or action potentials exhibit high levels of variability or stochasticity [21, 19] that have presented numerous challenges for scientists in uncovering how the brain functions. In particular, the random behavior of neurons in response to stimuli has resulted in the development of complicated theories for the role of stochasticity in neural processing [13]. Given that variability is an undeniable yet crucial aspect of cortical neural processing, it is not surprising that methods to characterize and analyze such behaviors have received so much attention in computational neuroscience.

In characterizing neural spiking activity, theorists often focus on the random time between spikes [19], although there are certainly many other entities that are of interest. This quantity is key to understanding the possible response of a given neuron. If the neuron resets completely after a spike, then the time between spikes is related to other quantities of interest (e.g., spike count in a given time window [7]). Although the reset or renewal assumption is not completely accurate, it has been successfully employed in theoretical investigations of coupled spiking neural networks [10] and in numerous experimentally driven models. The time series of the neural network activity undoubtedly contains more information than simply the time between spikes or *interspike interval* (**ISI**), but the ISI still contains a great deal of information and provides an in depth understanding to how likely a neuron is to respond at a given point in time. Thus, our study is focused on characterizing the probability distribution of the ISIs of a network of real neurons from *in vivo* recordings of awake adult macaque monkeys in area MT of the visual cortex [9].

In this paper, we employ a statistical framework based on parametric distribution fitting to characterize the random time between spikes using several goodness of fit criteria, including: maximum likelihood (**ML**),

and Akaike Information Criteria (**AIC**) [1, 4]. The data [9] contains the individual spike times of a population of neurons, where the awake animal is subject to complex motion where objects mimic realistic stimuli (see [8] for further details). The value of using parametric distributions over non-parametric or empirical based methods is in the necessity of developing sound theories of general neural processing. Specifically, modelers in computational neuroscience necessarily make assumptions for the statistics of neuron spiking behavior because mathematical analysis of such systems is only amenable with parametric descriptions (i.e., a common assumption is that neurons spike via a Poisson Process, corresponding to an exponential distribution of the ISI). In this study, we find that, despite all of the heterogeneity, only one family of parametric distributions was determined to be the best-fitting model for all neurons in the robust data set that contains a wide range of overall firing rates and variability. In the ISI probability density, we further test the distribution fitting methodology on a bursting model that contains multiple peaks and find the results still hold. Despite the wide variety of neural data and model data considered, we find that each of the goodness of fit criteria resulted in the *same* parametric distribution as the 'best' fit. Finally, we unravel the AIC formula to better understand why the same parametric distribution is consistently selected with AIC. Since ML increases with each additional parameter, that criteria consistently selected the distribution with the most parameters as the best-fitting distribution. This work sheds light on that nature of neural spiking statistics useful to theorists and modelers – although the parameters can vary widely, neural behavior can be captured by a particular parametric distribution depending on the goodness of fit criteria.

2   Materials and Methods

2.1   Parametric Distributions

We consider two commonly used parametric probability distribution functions (**PDF**) as the basis for the class of models considered in capturing the ISI data: the exponential distribution

$$f_{Exp}(x) = \frac{1}{\tau}e^{-x/\tau}, \text{ for } x > 0, \tag{1}$$

and the gamma distribution

$$f_{Gam}(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}, \text{ for } x > 0. \tag{2}$$

These two families of PDFs are by far the most commonly utilized distributions for the ISI PDF. A common framework in neural network modeling studies is to assume the 'background,' or indirect, synaptic inputs come from a stochastic spike train that is governed by a Poisson process, which is equivalent to having an exponential PDF f_{Exp} for the ISI density [17]. Furthermore, in more complicated frameworks where the spike train is governed by a general renewal process, the gamma PDF f_{Gam} is often assumed to be the underlying ISI density [12, 5, 10].

With these two base PDFs, we also consider the three possible mixture distributions as candidate models for a total of five candidate models:

$$f_{EE}(x) = c(\frac{1}{\tau_1}e^{-x/\tau_1}) + (1-c)(\frac{1}{\tau_2}e^{-x/\tau_2}), \text{ for } x > 0. \tag{3}$$

$$f_{EG}(x) = c(\frac{1}{\tau}e^{-x/\tau}) + (1-c)(\frac{1}{\beta^\alpha \Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}), \text{ for } x > 0. \tag{4}$$

$$f_{GG}(x) = c(\frac{1}{\beta_1^{\alpha_1}\Gamma(\alpha_1)}x^{\alpha_1-1}e^{-\frac{x}{\beta_1}}) + (1-c)(\frac{1}{\beta_2^{\alpha_2}\Gamma(\alpha_2)}x^{\alpha_2-1}e^{-\frac{x}{\beta_2}}), \text{ for } x > 0. \tag{5}$$

Equation (3) will be referred to as the Exponential-Exponential mixture distribution, equation (4) will be referred to as the Exponential-Gamma mixture distribution, and equation (5) will be referred to as the Gamma-Gamma mixture distribution. In a mixture distribution, the parameter c has a value between zero and one resulting in a weighted sum of two distributions. The other parameters also have constraints; table

Table 1: Constraints for parameters in parametric distributions (1)–(5)

Parameter	$\tau, \tau_{1/2}$	$\alpha, \alpha_{1/2}$	$\beta, \beta_{1/2}$	c
Constraint:	> 0	> 0	> 0	$\in [0, 1]$

1 contains all of the constraints. We remark that these five parametric PDFs encapsulate a wide range of behavior. Appendix A provides justification for using the unimodal exponential and gamma PDFs because the data appears to be unimodal. The three additional mixture distributions are able to capture multimodal PDFs (Fig. 6).

2.2 Data and Data Filtering

The data, which the statistical framework was applied to, was obtained from the freely available Pack Lab [9]. This data was obtained by presenting the awake animal with complicated visual stimuli (moving dots) in an effort to mimic realistic visual scenes.

Upon examination of the data obtained from *in vivo* awake rhesus macaque monkeys, we found exact duplicates that were discarded in order to have a better picture of the population of neurons (duplicate spike trains are presumably data processing/saving errors). The original data set consisted of recordings from 84 neurons, but we determined that there were three groups of duplicates, neurons 64-70, 72-76, 77-81, and 82-84 (their labeling). Thus, neurons 65-70, 73-76, 78-81, 83-84 were removed to obtain the largest possible unique set of neural recordings. In total, 16 of the 84 neural recordings were discarded to yield 68 unique neuron recordings. The total number of spikes from a given neuron varied widely, but had at least 1,276 spikes and up to 215,915 – with an average of 10s of thousands of spikes.

Note also that the overall average firing rates, defined as: $\dfrac{\text{Total \# Spikes}}{\text{Total time}}$, is in the physiological range of less than one Hz (little spiking) to 10's of Hz (Fig. 1a), with one neuron in particular having a larger rate than the others (neuron 68 has a firing rate of 199.75 Hz). A firing rate of 199.75 Hz is not uncommon for fast-spiking inhibitory neurons. The coefficient of variation (**CV**) is defined as the standard deviation (σ) divided by the mean [17, 19, 21]:

$$CV := \frac{\sigma}{\mu}. \tag{6}$$

The CV of the ISI is a common way to quantify the variability of a spike train [10]. In Figure 1b, we see that the dataset has a wide range of CV values that are generally larger than one (minimum CV is 0.94), with some neurons having CV values greater than four (very high levels of variability). These CV values are consistent with electrophysiological recordings from many experimental laboratories and in many brain areas [19, 6]. These preliminary observations lead to the conclusion that this dataset is a good representation of coupled cortical neural networks, with commonly observed firing rates and CV.

2.3 Goodness of Fit and MLE Fitting Procedure Details for Neural Recordings

Maximum likelihood estimation (**MLE**) from Matlab routine **mle.m** was used to calculate the parameters for the parametric distributions. After the MLE algorithm was applied, two main goodness of fit criteria were considered: maximum likelihood (ML) and the Akaike information criterion (AIC). The Bayesian information criterion (BIC) was also applied, but was not analyzed in further detail because it always yielded the same results as AIC and the distributions were not as close-fitting as AIC (Fig. 3). The precise definitions of these values are as follows.

Maximum likelihood (ML). For each neuron, the best model is the one with the **largest** log-likelihood

(a) The firing rate of recorded neurons.

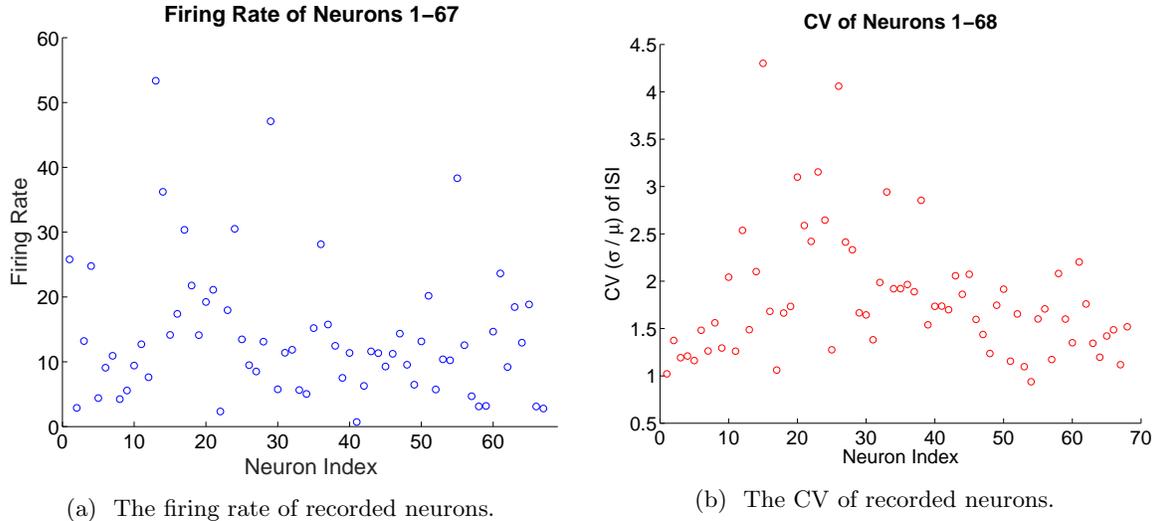(b) The CV of recorded neurons.

Figure 1: The population of neurons have spiking statistics consistent with other neural networks. (a) The firing rate of neurons 1-67 are shown above. Neuron 68 was not included because it had a firing rate of 199.75 Hz and would have obscured the axis scale. (b) The coefficient of variation (CV): standard deviation of the ISI divided by the mean, for all 68 neurons. The plot shows a wide variety of values, almost all of which are larger than 1, consistent with a lot of measurements of cortical neural activity.

value:

$$\log(L(\hat{\theta})) := \sum_{j=1}^{n} \log(f(t_j|\hat{\theta})\,\Delta t) \tag{7}$$

where t_j is an ISI data point and n is the number of data points, and Δt is the width of the bin, which we set to two milliseconds. For tractability, we assumed the ISI data points are independent, which is reasonable considering the sheer number of data points (see section 2.2). This formula naturally depends on the underlying parametric PDFs. For example, with the exponential distribution, $\log(L(\hat{\theta})) = \sum_{j=1}^{n} \log(\frac{1}{\tau}e^{-\frac{t_j}{\tau}})\Delta t$.

The finite size correction AIC [4]. This penalizes models with more parameters rather than simply considering the (log) likelihood. For each neuron, the best model is the one with the **smallest** AIC value:

$$\text{AIC} := -2\frac{\log(L(\hat{\theta}))}{n} + 2k + \frac{2k(k+1)}{n-k-1} \tag{8}$$

where k is the number of parameters in the model (e.g., $k = 5$ for f_{GG} in equation (5)).

The finite size correction BIC [18, 23]. This is similar to AIC; the best model is the one with the **smallest** BIC value:

$$\text{BIC} := -2\frac{\log(L(\hat{\theta}))}{n} + k\log(n). \tag{9}$$

The data from Chris Pack's lab (see section 2.2) we used consisted of recordings from 68 neurons, during which the animal was subject to complex visual motion mimicking real stimuli. Initially, we fit all five parametric distributions (see equations 1–5) using MLE to estimate the parameters. However, further investigation led to the realization that the MLE yielded $\alpha < 1$ for ISI data had values of zero for some neurons. For our statistical fitting methodology, this is problematic because the corresponding (log) likelihood values would be fraught with numerical errors (large numbers). These zero ISI values are directly obtained from the data [9] and were not introduced by finite resolution of our programs/methods. Zero ISI values are not physiological and were introduced via rounding errors because the spike times in the data had 1 ms resolution, so we decided to discard these points. Note that this is only a potential issue for three of the
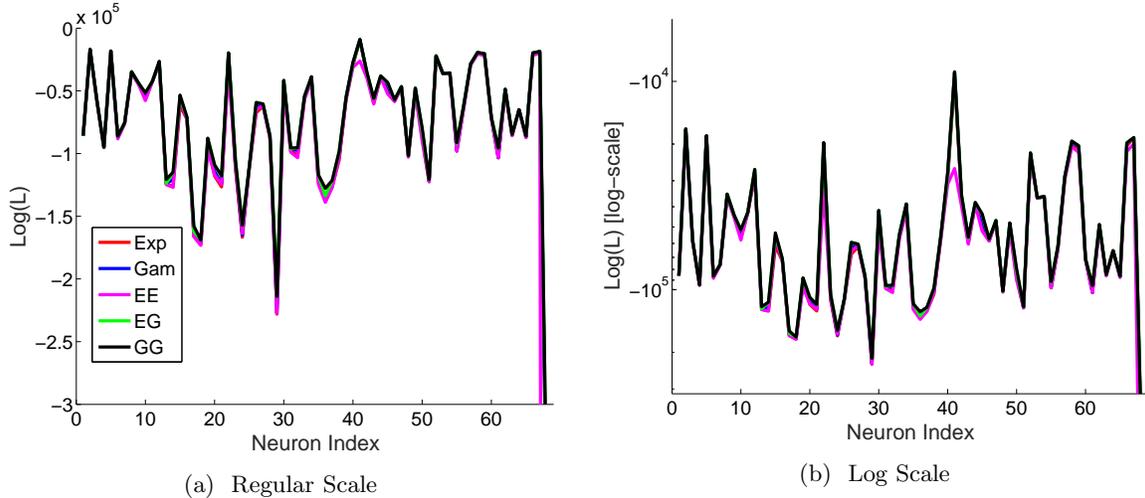
(a) Regular Scale   (b) Log Scale

Figure 2: Maximum Likelihood Goodness-of-Fit Criteria. A higher value indicates a better fit when using the ML criteria. For this dataset, although it is close for some neurons, the Gamma-Gamma mixture distribution f_{GG} (see equation (5)) is consistently above **all** of the others. Thus, it has the best fit for every neuron according to maximum likelihood criteria, where each plot represents the following. (a) The log likelihood for each neuron. (b) The same as (a) but the vertical axis is on a log-scale.

five parametric distributions considered: the Gamma distribution (equation (2)), the Exponential-Gamma mixture distribution (equation (4)), and the Gamma-Gamma mixture distribution (equation (5)). Also note that not all neurons (22 out of 68) had zero ISI values discarded, and when discarded these spike times were at most 0.75 percent of the total number of spike times for a given neuron. Thus, we are confident that this small amount of discarded data ($< 0.75\%$) does not effect our results. After this initial 'first pass' of fitting the five parametric PDFs to the data, we refit each of the five distributions to the augmented data to arrive at the final fit to the parameters.

3   Results

The previously described MLE of the parameters (see section 2.3) was used on the five parametric distributions to determine which of the five possible models 'best' describes the stochastic spiking activity of a particular neuron. For a more complete picture, we use three goodness of fit criteria initially.

When using the ML criteria on each of the five distributions, where the largest log-likelihood value indicates the best-fitting distribution, we found that the Gamma-Gamma mixture distribution (equation (5)) was the best fit for **all** of the neurons that were recorded, as shown in Figure 2. It may be hard to see in the plot, but taking the average of all 68 neurons gives the following results for each log likelihood. For Exponential, the average is approximately -0.000792. For Gamma, the average is approximately -0.000784. For the Exponential-Exponential mixture, the average is approximately -0.0000113. For the Exponential-Gamma mixture, the average is approximately -0.000759. And finally, it is the highest for the Gamma-Gamma mixture distribution with an average of -0.000756. This is not too surprising because among the 5 distributions, we would expect the Gamma-Gamma mixture to have the highest (log) likelihood value because it has the most parameters. Although it makes intuitive sense that adding more parameters likely increases the (log) likelihood values (especially with these parametric distributions composed of similar PDFs), from Figure 2, it is evident that the Gamma-Gamma mixture has pretty similar (log) likelihood values to the parametric distributions for some neurons. Thus, we use another goodness of fit criteria that penalizes extra parameters (AIC).
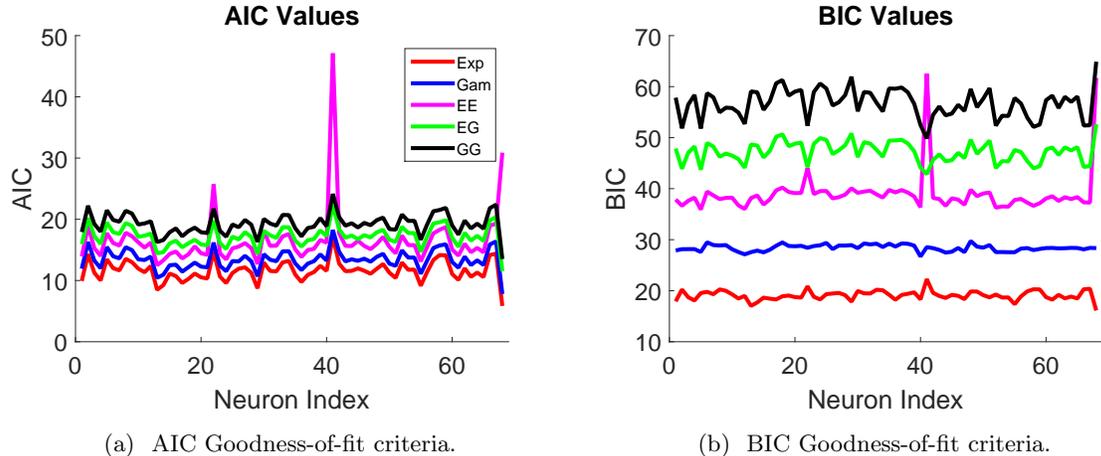
(a) AIC Goodness-of-fit criteria.	(b) BIC Goodness-of-fit criteria.

Figure 3: For both AIC and BIC, a lower value indicates a better fit. As shown, the exponential distribution f_{exp} (see equation (1)) has the best fit for **every** neuron according to both AIC and BIC criteria. (a) AIC values. (b) BIC values. We focus on AIC because the best distribution is less confounding with BIC than with AIC.

When using the AIC criteria, we found that the Exponential distribution (equation (1)) was the best fit for all of the neural recordings, as shown in Figure 3. When using the BIC criteria, the results were the same (the Exponential distribution was deemed best), but differences were magnified when compared to AIC (see Fig. 3). That is, the Exponential distribution is relatively better than the other distributions with BIC than with AIC. For this reason, we chose to focus on comparing just AIC and ML. We remain agnostic about which goodness-of-fit procedure to use because this likely depends on the purpose of the fitting and the model. However, our work will focus on why the AIC consistently selects the simple Exponential distribution as the best fitting.

Plotting the parametric ISI distributions for all 68 neurons and comparing to the empirical PDF (neural recordings) is obviously not feasible, so we focus on the best and worst fits for both AIC and ML criteria. The empirical PDF for each neuron was calculated by discretizing the ISI state space (time) into $\Delta t = 2\,\mathrm{ms}$ bins ranging from the minimum ISI to the maximum ISI value, and normalizing so the PDF integrates to 1. A plot of the four fits is provided in Figure 4. Notice how poorly the exponential distribution captured the smaller ISI values for neuron 41 (Plot 4a) despite being chosen as the best fitting according to AIC. Neuron 68 (Plot 4b) had parameters that yielded very similar distributions among all five considered distributions.

With these preliminary results, we were interested in investigating the following questions:

- How robust are these results? (AIC→Exponential, ML→Gamma-Gamma).

- Why does AIC, a ubiquitous goodness-of-fit criteria, always result in selecting the Exponential distribution, even when it does not capture most of the spiking statistics (smaller ISI values)?

We address these two questions in the next two sub-sections. The intuitive understanding of why the ML always yields the distribution with more parameters (Gamma-Gamma) does not merit further investigation.

3.1 Testing preliminary results on a bursting model neuron

To test the robustness of the statistical fitting results beyond a single data set, we considered a spiking neuron model (see Appendix B for details). This model has multiple state variables, time-scales, and is able to **burst** (i.e, fire many action potentials in a short period of time). Bursting is a real attribute in some neurons, and can result in a multimodal ISI PDF; here, we select parameters so that the ISI PDF is multimodal (see Fig. 8b). As described in Appendix A, the data, although real, is not multimodal. The
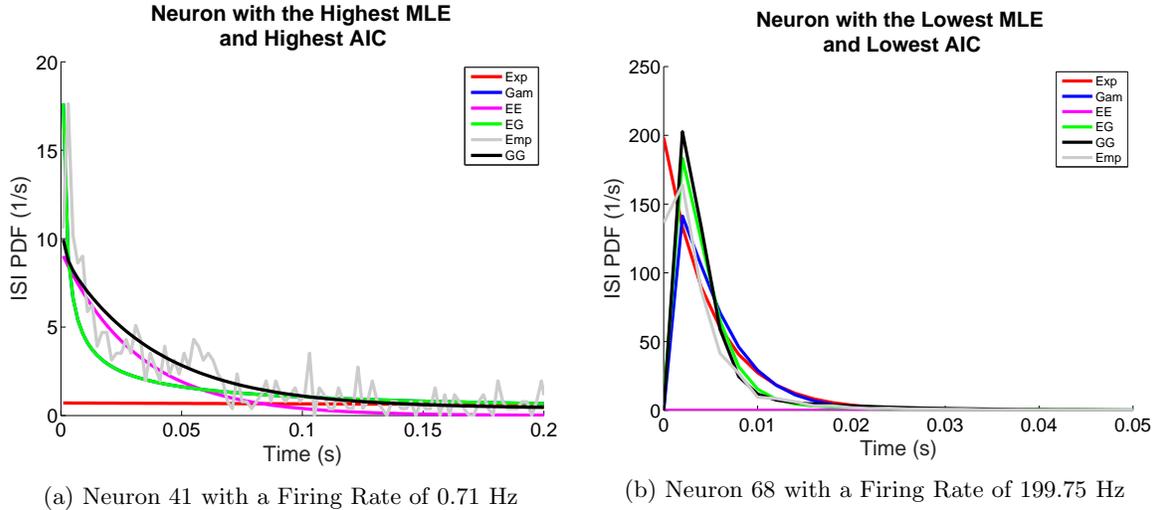
(a) Neuron 41 with a Firing Rate of 0.71 Hz

(b) Neuron 68 with a Firing Rate of 199.75 Hz

Figure 4: ISI PDFs and corresponding fits for the extreme (best and worst) cases using ML and AIC. With ML, the Gamma-Gamma mixture distribution f_{GG} (see equation (5)) is always the better fit, so the highest (best) likelihood is shown in a), corresponding to neuron 41, and the lowest likelihood (worst) is shown in b), corresponding to neuron 68. With AIC, the exponential distribution f_{exp} (see equation (1)) is always the better fit, so the lowest AIC (best) is shown in b), corresponding to neuron 68 and the highest AIC (worst) is shown in a), corresponding to neuron 41.

bursting model has a more complicated ISI PDF. The purpose of further testing is especially important because it might result in AIC giving another distribution besides the Exponential. The more complicated shape may give the more complicated parametric distributions a better chance.

Furthermore, to test how the size of the data might possibly affect the fitting results, we chose to vary the number of ISIs from the bursting neuron model to capture the range of the Pack Lab data. Specifically, we considered three numbers of ISI: i) 1,600 spikes is about the minimum ISI count in the Pack Lab data and is a small enough number that might effect the finite-correction AIC, ii) 10,000 spikes is within the range of the minimum and maximum ISI in the Pack Lab data, and iii) 20,000 spikes is on the higher end for the ISI count in the Pack Lab. As evident from Figure 5, the results are exactly as in the case with real neural recordings. Specifically, we found that the Exponential distribution fit the best according to AIC and the Gamma-Gamma mixture distribution fit the best according to maximum likelihood, even when using three vastly different numbers of ISIs. The empirical and fitted ISI PDFs for each of the different numbers of ISI are shown in Figure 6. Based on visual inspection, the Gamma-Gamma mixture distribution appears to be the best fitting distribution despite AIC choosing the Exponential distribution as the 'better' distribution. As shown in Figure 5, according to maximum likelihood criteria, the Gamma-Gamma mixture distribution has the highest (log) likelihood value while the remaining four are on top of each other. Both AIC and BIC are also included in Figure 5 and the three show the consistency with the data obtained from the macaque monkeys. Although the results would almost certainly not be applicable to all spike data from neurons, these results taken together are useful for modelers/theoreticians who use a single distribution (perhaps with different parameters) to describe the spiking activity of a variety of neurons.

3.2 The AIC penalty for more parameters is severe

Although the Exponential distribution was able to fit some of the data, there were some clear examples where it was not able to properly capture neural recordings and the spikes from the bursting neuron model, while maximum likelihood consistently resulted in a better fitting distribution (Gamma-Gamma, away from the tail of the PDF) for both the empirical data and the burst models. This motivated further investigation
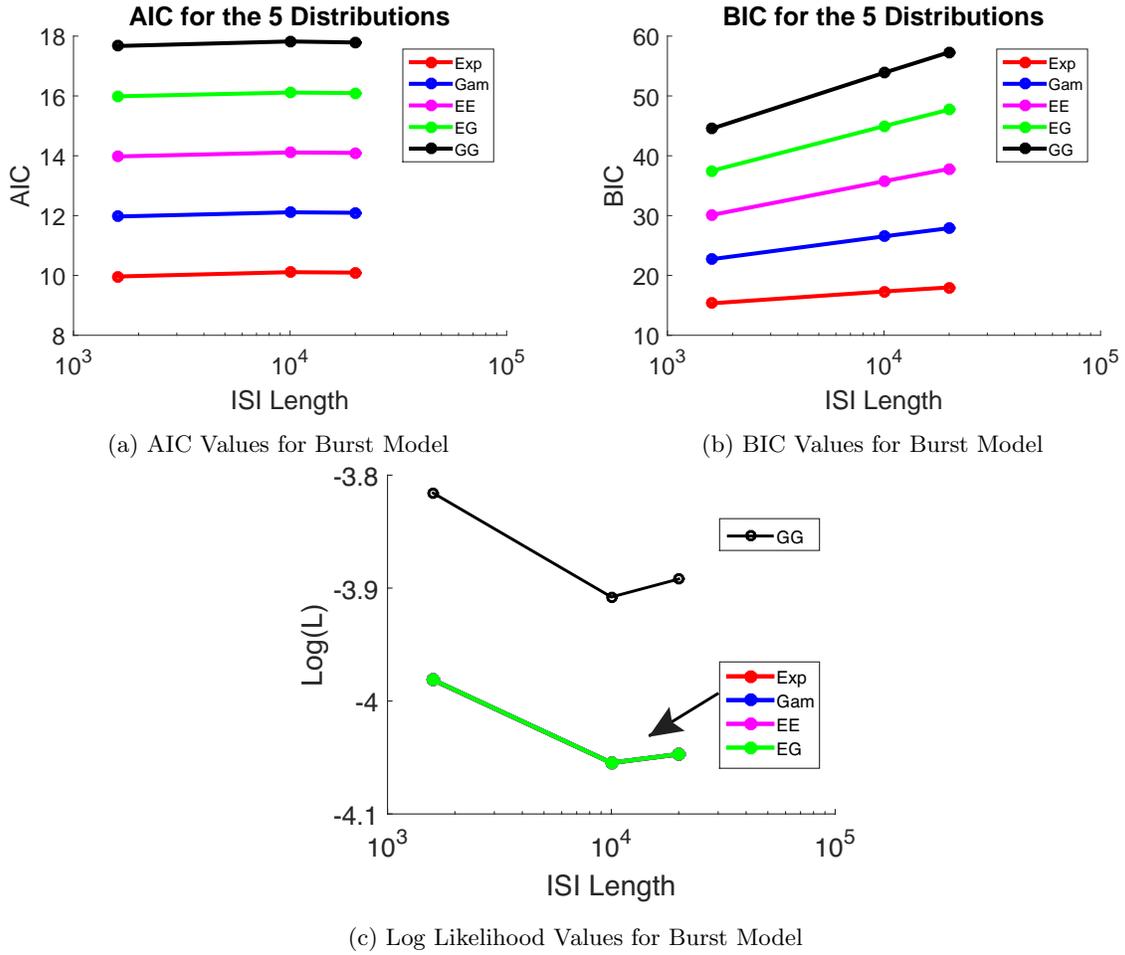
(a) AIC Values for Burst Model

(b) BIC Values for Burst Model

(c) Log Likelihood Values for Burst Model

Figure 5: The burst model gave the same results as with neural data. According to both AIC and BIC, the Exponential distribution f_{exp} was the best fit as shown in plots (a) and (b), respectively. Plot (c) shows that the Gamma-Gamma mixture distribution f_{GG} had the highest log likelihood according to maximum likelihood criteria. Note that in (c), the four remaining distributions are on top of each other so only the Exponential-Gamma mixture distribution f_{EG} is showing.

(a) Burst model with 1,600 spikes

(b) Burst model with 10,000 spikes

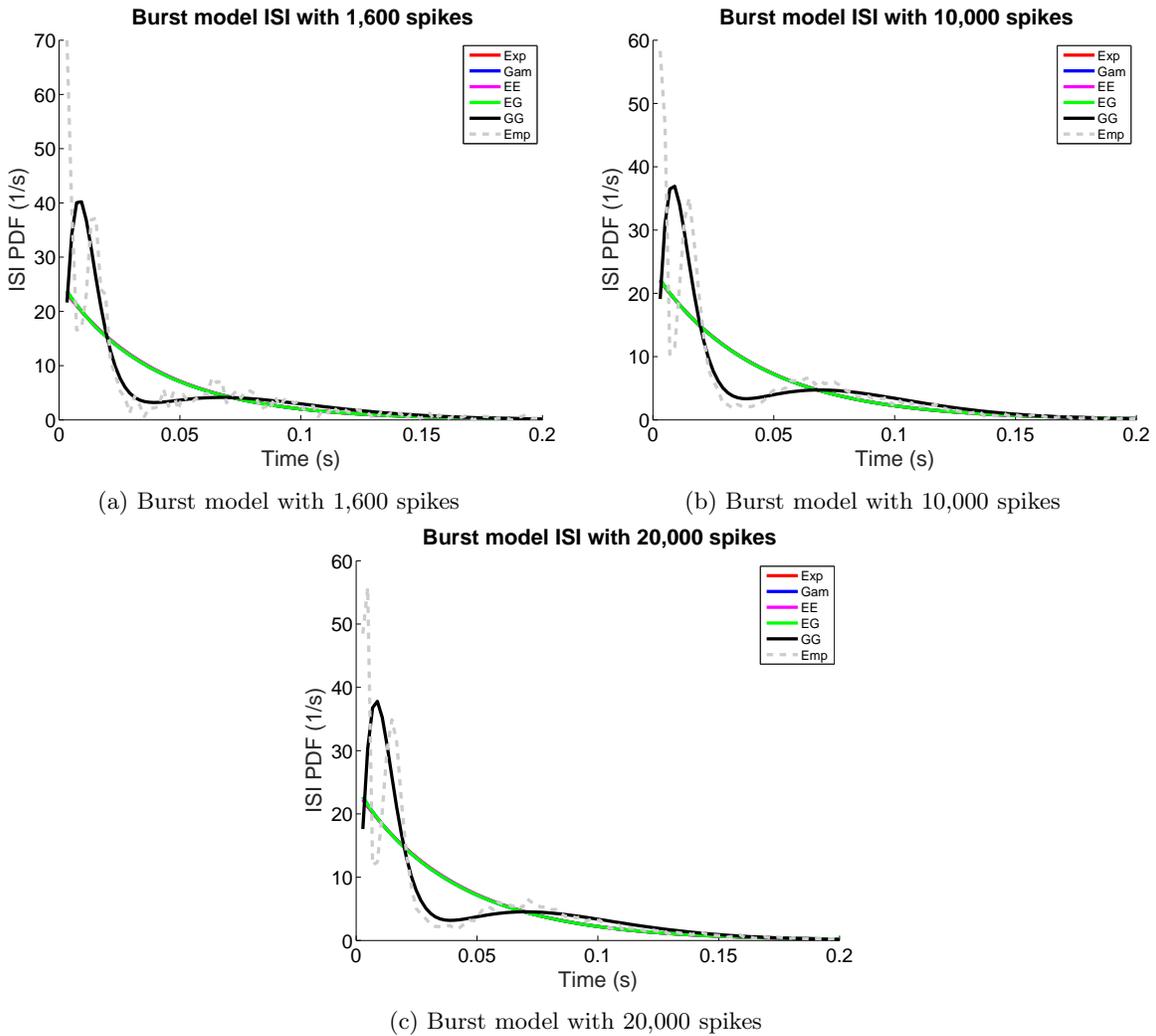(c) Burst model with 20,000 spikes

Figure 6: Clearly, visual inspection shows that the Gamma-Gamma mixture distribution f_{GG} (black line) is the best fit, indicating that ML is a better determination of best fit. However, the tail may imply different results, indicating that AIC could be better. Note that the four remaining distributions are on top of each other so only the Exponential-Gamma mixture distribution f_{EG} is visible.

of the AIC formula.

To better understand the AIC, consider comparing two models only, where one model has k parameters and the other has one more parameter, $(k+1)$. Let model A represent a distribution with k parameters and model B represent a distribution with $(k+1)$ parameters. Furthermore, consider the AIC (Eq. (8)) without the last term $\frac{2k(k+1)}{n-k-1}$. Assuming that n is much greater than k, the last term can be ignored. We have:

$$Model \quad A : AIC = \frac{-2l_1}{n} + 2k, \quad \text{where} \quad l_1 = \sum_{j=1}^{n} \log(\rho_1(x_j|\theta_1))$$

$$Model \quad B : AIC = \frac{-2l_2}{n} + 2(k+1), \quad \text{where} \quad l_2 = \sum_{j=1}^{n} \log(\rho_2(x_j|\theta_2)).$$

The variable θ_i represents the parameters of the model. In this study,

$$\theta_i = \begin{cases} \tau \\ \alpha, \beta \\ \tau_1, \tau_2, c \\ \tau, \alpha, \beta, c \\ \alpha_1, \beta_1, \alpha_2, \beta_2, c \end{cases}$$

In order for Model B to have a lower AIC and be chosen as a better model over Model A, we need

$$\frac{-2l_2}{n} + 2(k+1) < \frac{-2l_1}{n} + 2k.$$

This is equivalent to:

$$\frac{-l_2}{n} + \frac{l_1}{n} < -1 \quad \Longleftrightarrow \quad l_2 - l_1 > n$$

Therefore, the extra parameter in Model B needs to increase the log likelihood by one on average. Exponentiating both sides yields:

$$e^{l_2 - l_1} > e^n,$$

where

$$e^{l_2 - l_1} = \frac{\prod_{j=1}^{n} \rho_2(\theta_2)}{\prod_{j=1}^{n} \rho_1(\theta_1)}.$$

Thus, in adding an extra parameter, the ratio of the likelihoods has to increase by a **huge** number (e^n) for Model B to be selected.

Here we have shown that adding a parameter can make that distribution extremely difficult to be chosen according to AIC criteria, at least in the context of this ISI data. When the full AIC equation (Eq. (8)) is used, the last equation is as follows:

$$e^{l_2 - l_1} > e^n e^{n[\frac{(k+1)(k+2)}{n-k-2} - \frac{k(k+1)}{n-k-1}]}.$$

Note that the exponent in the second factor is positive, so that $e^{n[\frac{(k+1)(k+2)}{n-k-2} - \frac{k(k+1)}{n-k-1}]} > 1$, thus including the last (finite-size correction term) would make the right-hand side of the last inequality an even larger number. This is clearly a huge number and would be difficult for any one distribution with more parameters to be selected, at least among the five PDFs we considered.

4 Discussion

We applied statistical methods and various goodness of fit criteria to spiking neural data obtained from an awake animal to better understand the variability of neural activity in a representative population of neurons. Using a formal statistical framework to characterize spiking neural data is under-utilized by computational neuroscientists, who often use a single goodness-of-fit criteria (often ML). Our work gives a better understanding for how different goodness-of-fit criteria yield different results.

The dataset contained several aspects of coupled cortical neural networks making it a good candidate for this study: (i) reasonable firing rates with a wide range, likely including excitatory neurons (low rates) and inhibitory neurons (higher rates); (ii) the CV of the ISI around one or above; (iii) the CV of the ISI have a wide range. Yet despite all of this realism, we found that only one family of parametric distributions was selected as the best for all goodness-of-fit criteria (although the the distribution depended on which criteria). According to ML criterion, the Gamma-Gamma mixture distribution consistently had the best fit and according to AIC and BIC criterion, the Exponential distribution consistently had the best fit. Since this dataset contains many generic features of cortical neural networks, some of these results may have wider applicability beyond these 68 recorded neurons. Indeed, we applied the statistical framework to a bursting spiking model with a multimodal ISI PDF, and found that the results still hold.

The results here should be valuable to modelers in computational neuroscience in regards to the necessary simplifying assumptions of neural spiking behavior (often assuming a parametric distribution for the ISI) made in order for specific systems to be amenable to mathematical analysis. In this vein, our study sheds light on how the underlying assumption of a spike train being a Poisson process (exponential ISI distribution) or not. It is important to note that the Poisson process assumption for a given spike train, and issues surrounding the validity of this assumption, are contemporary issues that are so ubiquitous that we only provide a small subset of references [22, 12, 5, 3, 11, 15, 20].

Acknowledgments

Appendix A: Evidence that ISI data is unimodal.

We performed an analysis of the Pack Lab data using a freely available Matlab function `peakfinder.m` [24], which takes the ISI data and other inputs and returns the number of peaks. We used three inputs with the following values: `thresh=0`, `extrema=1`, and `sel=3/7*`$\left(\max f_{edf} - \min f_{edf} \right)$, where f_{edf} is the empirical distribution function formed by binning the ISI data into $2\,\text{ms}$ sized bins. `thresh` is the absolute threshold that the pdf has to be greater than in order for the function to return an extrema (set to zero since the empirical distribution function f_{edf} are all positive); `extrema` identifies the type of extrema (1 corresponds to a local maxima, -1 corresponds to a local minima); `sel` determines the amount of data that must be surrounding a point for it to be considered a peak (larger values make it more selective in determining whether or not each point is a peak).

The result of using this function on all 68 neuron's ISI data returned a single peak for all but five neurons. The five neurons that have multiple peaks are shown in Figure 7. By visual inspection we see that the empirical distribution functions arguably have one dominant peak despite the results of the `peakfinder.m` function, and the "multiple" peaks are quite noisy. Overall, these observations provide justification for using the unimodal exponential and gamma distribution functions as the base models, along with pairs of mixture distributions, to capture the ISI data.

(a) Neuron 2

(b) Neuron 5
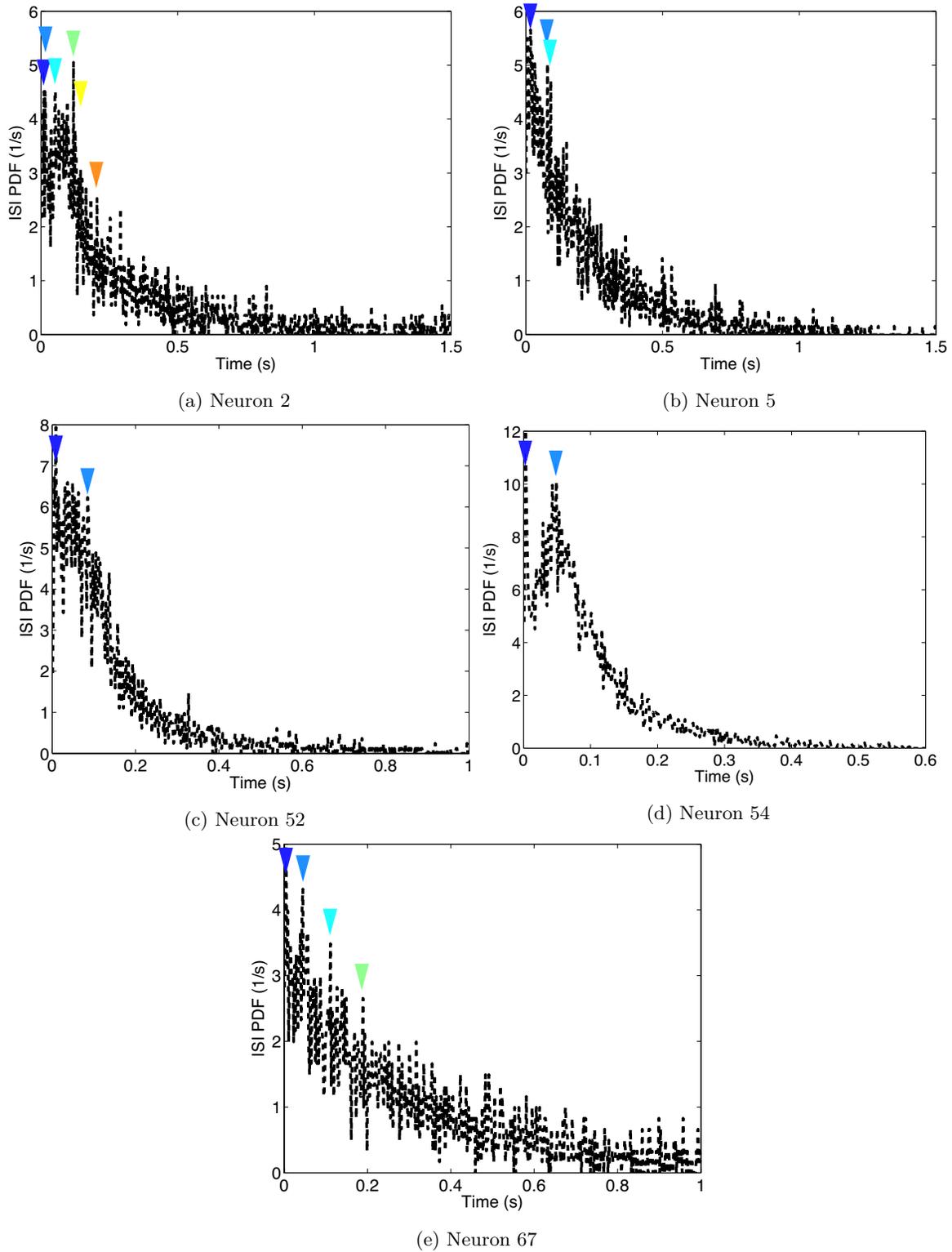
(c) Neuron 52

(d) Neuron 54

(e) Neuron 67

Figure 7: Neurons from the Pack Lab [9] where more than one peak was detected with `peakfinder.m`, indicated by colored arrows.

Table 2: Parameters for bursting model

Parameter	τ_m	I_{dc}	α	ref_s	β	γ	τ_η	τ_b	μ_1	μ_2	μ_3	μ_4
Values:	10 ms	0.8	20	0.8 ms	0.1	0.008	5 ms	10 ms	0.5	2	0.0007	0.025

Appendix B: Model of bursting neuron with multi-modal ISI distribution

The model used to exhibit random bursts of action potentials is based on the standard leaky integrate-and-fire neuron model, but with a depolarizing-after-potential (DAP) that is the burst mechanism. This model was chosen because it is simple, fast to simulate, does not directly rely on inputs from other neurons to induce bursting, and has been used by others for a relatively long period of time [16, 2, 14].

The equations for the model are:

$$\tau_m \frac{dv}{dt} = -v + I_{dc} + \eta(t) + \alpha \cdot DAP(t) \tag{10}$$

$$v(t^*) \geq 1 \text{ spike recorded}; \Rightarrow v(t^* + ref_s) = 0$$

$$DAP(t) = \begin{cases} z(t - t^*, \beta b(t^*)) - z(t - t^*, \gamma), \text{ if } t - t^* > ref_s \text{ and } t - t^* > ref_d(t^*) \\ 0, \text{ otherwise} \end{cases} \tag{11}$$

$$\tau_\eta \frac{d\eta}{dt} = -\eta + \sqrt{\tau_\eta}\xi(t) \tag{12}$$

$$\tau_b \frac{db}{dt} = -b + \tau_b \left(\mu_1 + \mu_2 b^2\right) \sum_k \delta(t - t_k^*) \tag{13}$$

$$ref_d(t) = \mu_3 + \mu_4 b(t^*) \tag{14}$$

$$z(t, a) = \frac{t}{a} e^{-t/a} \tag{15}$$

where t_k^* is the time of the last spike (recorded when $v(t) \geq 1$), and t_k^* is the time of the kth spike, $\xi(t)$ is a white noise process with zero mean and unit variance. The parameters are shown in table 2.

Figure 8a shows an example voltage trajectory of the bursting neuron in blue, with the black stars on top representing the spike times. This plot illustrates random behavior, with periods of rapid firing (bursts) as well as intermittent firing. The ISI density is indicative of this behavior (see Fig. 8b), as evidenced by the multiple peaks.

References

[1] H. AKAIKE, *Information theory and an extension of the maximum likelihood principle*, in Selected Papers of Hirotugu Akaike, Springer, 1998, pp. 199–213.

[2] K. BOL, G. MARSAT, J. MEJÍAS, L. MALER, AND A. LONGTIN, *Modeling cancelation of periodic inputs with burst-stdp and feedback*, Neural Networks, 47 (2013), pp. 120–133.

[3] A. BURKITT, *A review of the integrate-and-fire neuron model: Ii. inhomogeneous synaptic input and network properties*, Biological cybernetics, 95 (2006), pp. 97–112.

[4] K.P. BURNHAM AND D.R. ANDERSON, *Model selection and multimodel inference: a practical information-theoretic approach*, Springer Science & Business Media, 2002.

[5] H. CÂTEAU AND A. REYES, *Relation between single neuron and population spiking statistics and effects on network activity*, Physical review letters, 96 (2006), p. 058101.

(a) Voltage trajectory of burst model neuron
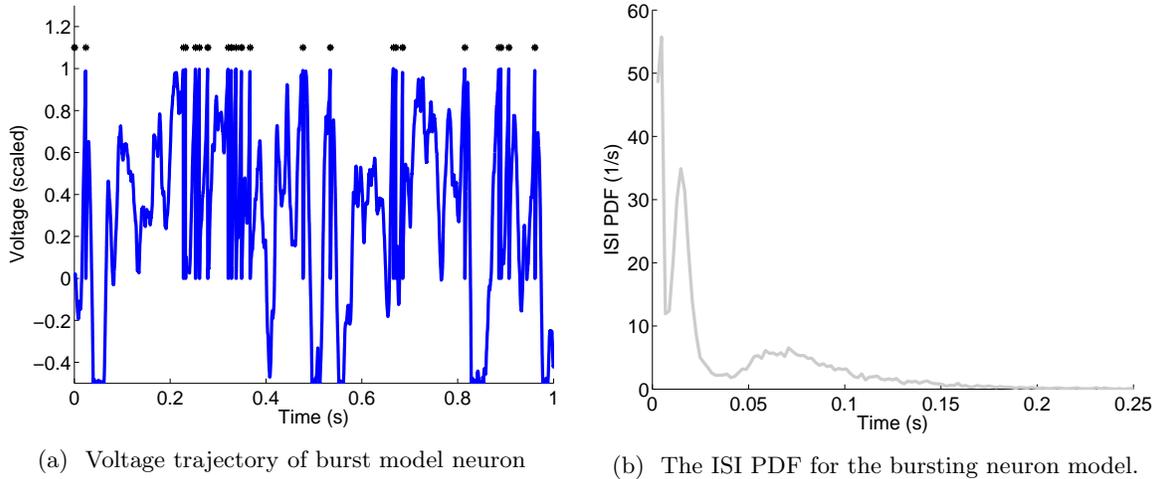
(b) The ISI PDF for the bursting neuron model.

Figure 8: a) An example voltage trajectory of the bursting neuron model (see equations (10)–(15)). The model neuron exhibits complicated stochastic behavior with periods of bursting (lots of spikes), and quieter periods with less spiking. b) The corresponding ISI probability density from model simulations with the first 20,000 spikes.

[6] M. CHURCHLAND, B. YU, J. CUNNINGHAM, L. SUGRUE, M. COHEN, G. CORRADO, W. NEWSOME, A. CLARK, P. HOSSEINI, B. SCOTT, ET AL., *Stimulus onset quenches neural variability: a widespread cortical phenomenon*, Nature neuroscience, 13 (2010), pp. 369–378.

[7] D.D.R. COX AND V. ISHAM, *Point processes*, vol. 12, CRC Press, 1980.

[8] Y. CUI, D.L. LIU, F.A. KHAWAIA, C.C. PACK, AND D.A. BUTTS, *Diverse suppressive influences in area mt and selectivity to complex motion features*, The Journal of neuroscience, 33 (2013), pp. 16715–16728.

[9] ———, *Spiking activity in area mt of awake adult macaques in response to complex motion features*, CRCNS.org, (2013).

[10] W. GERSTNER AND W. KISTLER, *Spiking Neuron Models*, Cambridge University Press, Cambridge, United Kingdom, 2002, ch. 5, pp. 147–163.

[11] R. KASS AND V. VENTURA, *A spike-train probability model*, Neural computation, 13 (2001), pp. 1713–1720.

[12] C. LY AND D. TRANCHINA, *Spike train statistics and dynamics with synaptic input from any renewal process: a population density approach*, Neural computation, 21 (2009), pp. 360–396.

[13] M. MCDONNELL AND L. WARD, *The benefits of noise in neural systems: bridging theory and experiment*, Nature Reviews Neuroscience, 12 (2011), pp. 415–426.

[14] J. MEJIAS, G. MARSAT, K. BOL, L. MALER, AND A. LONGTIN, *Learning contrast-invariant cancellation of redundant signals in neural systems*, PLoS Computational Biology, 9 (2013), p. e1003180.

[15] E. MULLER, L. BUESING, J. SCHEMMEL, AND K. MEIER, *Spike-frequency adapting neural ensembles: beyond mean adaptation and renewal theories*, Neural Computation, 19 (2007), pp. 2958–3010.

[16] L. NOONAN, B. DOIRON, C. LAING, A. LONGTIN, AND R. TURNER, *A dynamic dendritic refractory period regulates burst discharge in the electrosensory lobe of weakly electric fish*, The Journal of neuroscience, 23 (2003), pp. 1524–1534.

[17] S. Ross, *Probability Models*, Cengage, 2010.

[18] G. Schwarz et al., *Estimating the dimension of a model*, The annals of statistics, 6 (1978), pp. 461–464.

[19] M.N. Shadlen and W.T. Newsome, *The variable discharge of cortical neurons: implications for connectivity, computation, and information coding*, The Journal of neuroscience, 18 (1998), pp. 3870–3896.

[20] W. Softky and C. Koch, *Cortical cells should fire regularly, but do not*, Neural Computation, 4 (1992), pp. 643–646.

[21] W.R. Softky and C. Koch, *The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps*, The Journal of Neuroscience, 13 (1993), pp. 334–350.

[22] C. Stevens and A. Zador, *When is an integrate-and-fire neuron like a poisson neuron?*, Advances in neural information processing systems, (1996), pp. 103–109.

[23] E. Wit, E. van den Heuvel, and Jan-Willem Romeijn, *?all models are wrong...?: an introduction to model uncertainty*, Statistica Neerlandica, 66 (2012), pp. 217–236.

[24] N Yoder, *peakfinder()*, Mathworks.com, (2015), pp. www.mathworks.com/matlabcentral/fileexchange/25500–peakfinder–x0–sel–thresh–extrema–includeendpoints–interpolate–.