A simple mechanism for higher-order correlations in integrate-and-fire neurons

David A. Leen* and Eric Shea-Brown†
Department of Applied Mathematics, University of Washington.
(Dated: February 26, 2013)

Here we show that a population of exponential integrate-and-fire (EIF) neurons receiving common input cannot be well described by a pairwise maximum entropy model. Common input into the EIF population hence gives rise to higher-order correlations. A tractable reduction of the EIF model, the linear-nonlinear (LN) cascade model, gives an order of magnitude improvement over the PME model. The LN cascade model receiving common input also gives rise to higher-order correlations. The Dichotomized Gaussian (DG) model with the appropriate first and second order moments gives an exact description of the EIF model receiving common input. The LN cascade model can be viewed as an approximation to the DG model and explains the successes of the DG model in the literature.

- Check usage of covariance vs. correlation (in many places, should it be the latter?)
- In intro ... work in notion that these (noise) correlations come from some network mechanism. A typical one is common input. Is that mech consistent with pwise or HOC? Answer HOCC ...

I. INTRODUCTION

Interest in the collective dynamics of neural populations is exploding, as new recording technologies yield views into neural activity on larger and larger scales [**? ?**] and new statistical analyses yield potential consequences for the neural code [**? ? ? ?**]. A fundamental question that arises as we seek to quantify these population dynamics is the statistical *order* of interactions among spiking activity in different neurons. That is, can the co-dependence of spike events in a set of neurons be described by an (overlapping) set of correlations among pairs of neurons, or are there irreducible higher-order dependencies as well? Recent studies show that purely pairwise statistical models are successful in capturing the spike outputs of neural populations under some stimulus conditions [**? ? ?**], but that different populations or stimuli can produce beyond-pairwise interactions [**? ? ? ?**].

This paper seeks to elucidate one key mechanism that can determine the pairwise vs. higher-order extent of statistical interactions. This is common – or *correlated* – input fluctuations arriving simultaneously at multiple neurons [**? ? ?**]. Previous studies used a simple, step function thresholding mechanism to come to an important conclusion: common, gaussian input fluctuations, when "dichotomized" so that inputs over a given threshold produce spikes, produce strong beyond-pairwise correlations in the spike output of multiple cells [1, 2]. This is an interesting finding, as the thresholding mechanism produces higher-order correlations in spike outputs starting with purely pairwise (gaussian) inputs.

A natural question is whether more realistic, dynamical mechanisms of spike generation – beyond "static" step function transformations – will also serve to produce strong higher-order correlations based on common input processes.

An exponential integrate-and-fire population with common inputs: A ubiquitous situation in neural circuitry is a cell population receiving common input (citations from Yu paper). ... we model this via a homogeneous population of N exponential integrate-and-fire (EIF) neurons, receiving common white noise inputs $\xi_c(t)$ and independent white noise inputs $\xi_i(t)$. Each cell's membrane voltage evolves according to:

$$\begin{aligned} \tau_m V_i' &= -V_i + \psi(V_i) + I_i(t), \\ I_i(t) &= \gamma + \sqrt{\sigma^2 \tau_m}\left[\sqrt{1-\lambda}\xi_i(t) + \sqrt{\lambda}\xi_c(t)\right], \end{aligned} \tag{1}$$

where: $\psi(V_i) = \Delta_T \exp\left((V_i - V_S)/\Delta_T\right)$ for the EIF neuron [E: cite Fourcard-Troume and Abbott-Dayan book]. [E: Compress text / ref to caption of Fig 1 for parameters] Here, τ_m is the membrane time constant which we set as $\tau_m = 5$ms, Δ_T controls the slope of the action potential initiation, taken to be 3mV. We use the usual convention of declaring a spike whenever the voltage reaches a certain threshold V_T, which we have set at 20mV. The EIF neuron has an additional "soft" threshold, V_S beyond which the voltage diverges to infinity. We set this value as -53mV. After a spike the voltage is reset to the rest potential V_R, set at -60mV, and is held there for the duration of the refractory period τ_{ref} of 3ms.

The input current has a constant (DC) component γ, which sets the equilibrium (rest) potential, and a stochastic noise component with amplitude σ. The input is modeled by a correlated Gaussian with mean γ, and covariance λ. We tune the noise amplitude so that when the DC component of the input is set to be $\gamma = -60$mV, the neurons fire at 10Hz; this yields $\sigma = 6.23$mV.

*Electronic address: dleen@washington.edu
†Electronic address: etsb@washington.edu

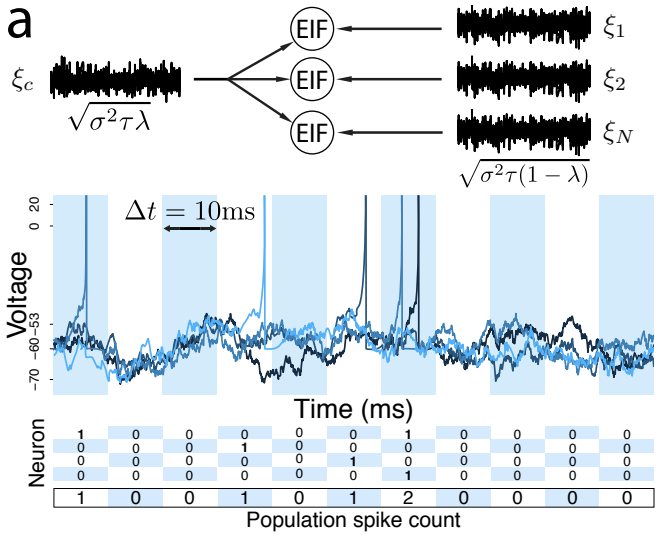

FIG. 1: (a) Exponential integrate-and-fire neurons receiving common input ξ_c and independent inputs ξ_i. The voltages of the neurons evolve according to equation [1]. Time is divided into bins of temporal resolution 10ms. A spike occurs when the voltage reaches a threshold of 20mV. The voltage is then reset to the resting potential -60mV and another spike cannot occur during the refractory period of 3ms. Spikes recorded from each of the EIF neurons in a bin contribute towards the population spike count. More than one spike occurring from the same neuron within a single bin is treated as a single event. This happens less than 0.4% of the time in our numerical simulations with $\mu = 0.1$ and $\rho = 0.1$.

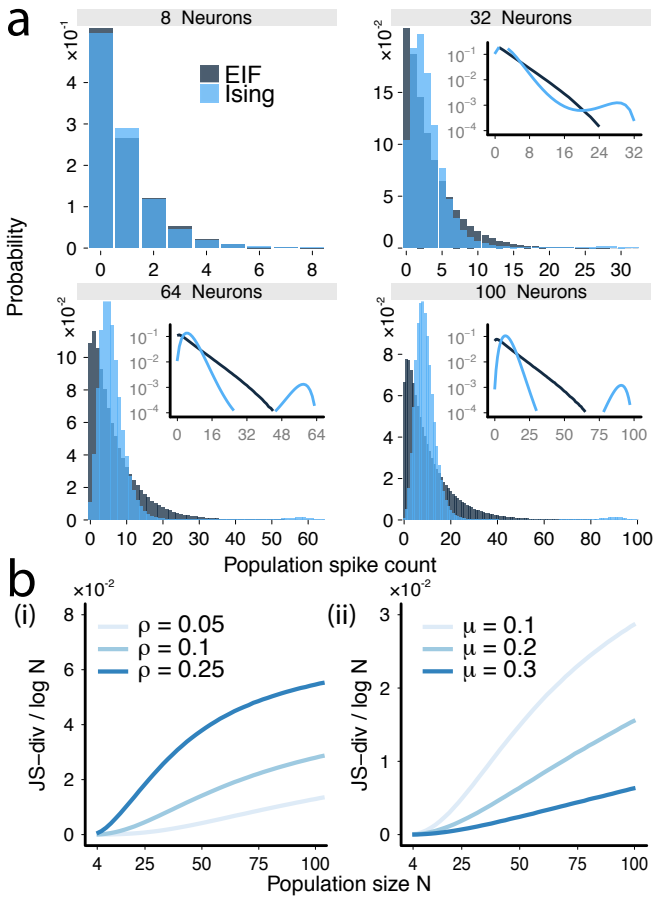

FIG. 2: (a) Population spike-count distributions for the EIF model and it's [E: no apostrophe!] second-order approximation for $8, 32, 64,$ and 100 neurons for $\mu = 0.1$ and $\rho = 0.1$. The distributions are similar for smaller populations and different for large populations. Inset: the same distributions on a log-linear scale. (b) The Jensen-Shannon (JS) divergence between the EIF and the pairwise maximum entropy (PME) model . We normalized by the $log(N)$, which is a natural growth rate of this measure. (Left) JS divergence for a constant value of $\mu = 0.1$ and different values of ρ, and (Right) for constant value of $\rho = 0.1$ and different values of μ as the population size increases. The JS-divergence grows with increasing correlation ρ and decreasing mean firing rate μ.

The spikes are binned into bins of temporal resolution T_{bin}, which we choose to be 10ms. On rare occasions ($< 0.4\%$ of bins, see Fig. 1 caption) multiple spikes from the same neuron can occur in the same bin. These are considered as a single spike. The output firing rate of is quantified by μ, the probability of a spike occurring in a bin. Pairwise correlations between the spiking in simultaneous bins of different neurons is quantified by the correlation coefficient $\rho = \alpha/\mu(1-\mu)$, where α is the corresponding covariance.

Emergence of strong higher-order correlations. As in [1**? ? ?**], we describe the network-wide firing statistics via the distribution of population spike counts (i.e., the number of simultaneously firing cells out of a maximum of N). Fig. 2a shows that this distribution has a strongly skewed shape. Do beyond-pairwise correlations plan an important role in determining this structure? To answer this, we compare the population spike-count distributions for the EIF model and its second-order approximation via a pairwise maximum entropy model: [E: ... give Ising formula here. I suggest writing it $P_{PME}(k) = ...$ in text. Could also state is equiv. to Ising...)] For small populations the PME does a satisfactory job. For populations larger than about $N = 30$ neurons the PME fails to capture the shape of the distribution. See figure 2a.

We quantify the difference between the PME and EIF population output distributions via the normalized JS divergence (see Fig. 2 (b)), for a range of population sizes N, firing rates μ, and correlation levels ρ. This demonstrates several important facts. The first is that the EIF model continues to produce strong beyond-pairwise correlations at a wide range of operating points. Additionally, as in [**? ?**], the Jensen-Shannon divergence grows with increasing population size N. Moreover, the divergence increases with pairwise correlation and decreases with firing rate over the ranges shown.

A semi-analytic spiking model that produces higher-order correlations. We next derive [E: emphasize analytical approach / cts model / investigate how the classi-

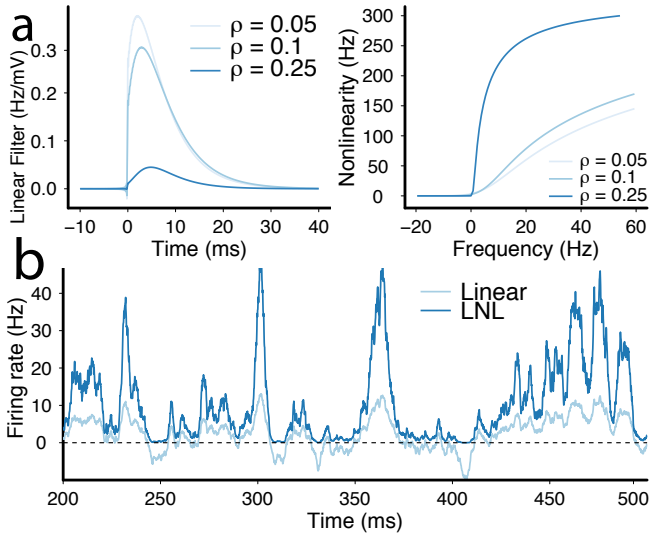

FIG. 3: (a) The linear filter $A(t)$ and static non-linearity for different values of the correlation coefficient ρ. The filter receives a noise amplitude of $\sigma\sqrt{1-\lambda}$. The static-nonlinearity receives a noise amplitude of σ. (c) The static non-linearity applied to the linear estimate of the firing rate, for $\mu = 0.1$, $\rho = 0.1$. The non-linearity increases the firing rate magnitude and rectifies negative firing rates. The firing rate is used to generate spikes, treating the spike output from repeated presentations of the firing rate as coming from a population of uncoupled neurons to arrive at a population spike count such as in (b). [E: In b, add a line for the "actual" firing rate of the EIF cells. As in Fig. 1 of ostojic paper]

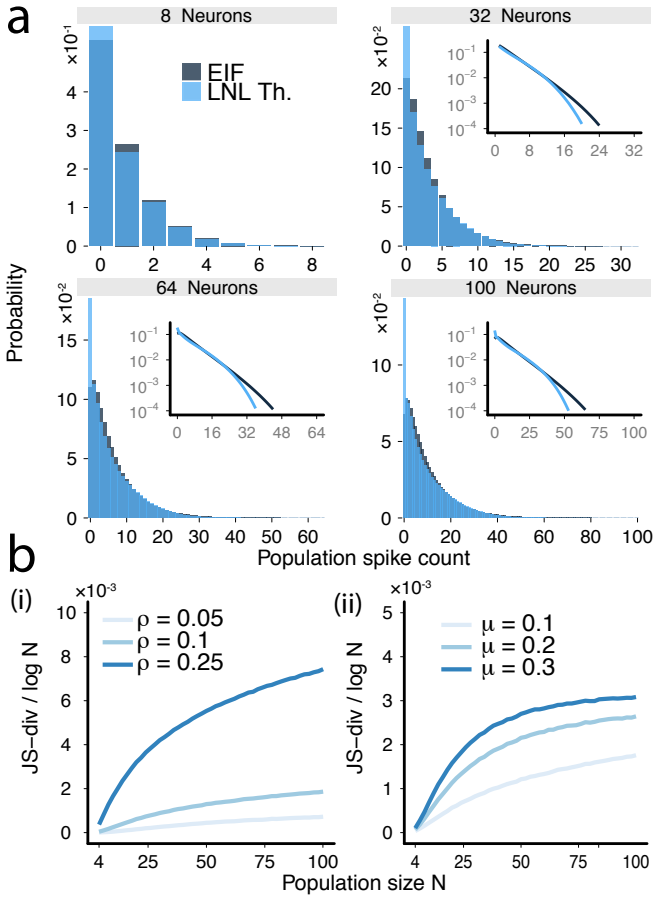

FIG. 4: The linear-nonlinear cascade gives a good approximation of the EIF population distributions. (a) A comparison between the population spike-count distributions for the EIF model and the linear-nonlinear cascade approximation for $8, 32, 64,$ and 100 neurons for $\mu = 0.1$ and $\rho = 0.1$. The LNL model greatly overestimates the zero population spike count probabilities. One reason is that there is no refractory period. The LNL model underestimates the tails of the probability distributions. This is because of the double counting, when truncating spikes in a bin the larger spike counts are penalized to a greater extent. Inset: the same distributions on a log-linear scale. (b) JS-divergence, order of magnitude smaller than PME, possibly converges to some limit? The order of the mean firing rates is reversed when compared to the PME because the LNL cascade gives a better approximation at higher firing rates μ, less problems with negative firing rates...

cal approach based on lin response theory works ...] an approximation for the EIF population statistics. This is a linear-nonlinear cascade, where each neuron fires as an inhomogeneous Poisson process with rate given by convolving a temporal filter $A(t)$ with an input signal $c(t)$ and then applying a time independent nonlinear function F:

$$r(t) = F\left(A * c(t)\right)$$

[Ostoijic]. The signal for each cell is the common input, so that $c(t) = \sqrt{\sigma^2\tau\lambda}\ \xi_c(t)$. $A(t)$ is computed as the linear response of the firing rate to a weak input signal, via an expansion of the Fokker Planck equation for Eqn. (1) around the equilibrium obtained with "background" current $\gamma + \sqrt{\sigma^2\tau(1-\lambda)}\ \xi(t)$. This calculation follows exactly the methods described in [Richardson]. For the static nonlinearity, we follow [Ostoijic] and take $F(x) = \Phi\left(\gamma + \frac{x}{\Phi'(\gamma)}\right)$, where $\Phi(\gamma)$ is the equilibrium firing rate obtained at the background currents described above. This choice, in particular, ensures that we recover the linear approximation $r(t) = A * c(t)$ for weak input signals.

[E: Insert some references to Fig. 3 above + below]

For an inhomogeneous Poisson process with rate $r(t)$ conditioned on a common input $c(t)$ the probability of at least one spike occurring in the interval $[t, t+\Delta t]$ is:

$$P(\text{spike} \in \Delta t|c) = 1 - \exp\left(-\int_0^{\Delta t} r(s+t)\mathrm{d}s\right).$$

Introducing the notation $\mathcal{S} = \int_0^{\Delta t} r(s)\mathrm{d}s$ we see that $P(\text{spike} \in \Delta t|c) = 1 - \exp(-\mathcal{S}) \equiv L(\mathcal{S})$.

Conditioned on the common input – or, equivalently,

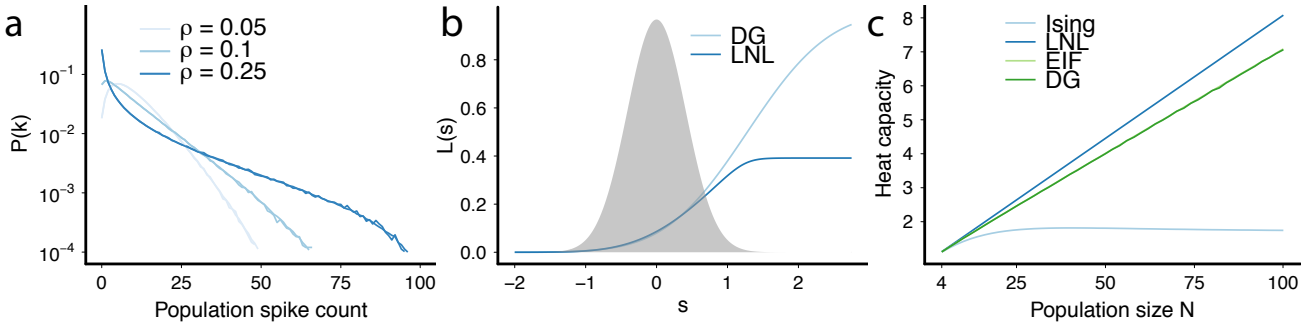

FIG. 5: (a) The Dichotomized Gaussian (DG) model gives an excellent description of the exponential integrate-and-fire (EIF) population spike count probability distributions across a range of correlation coefficient values. The two models are plotted on top of one another and appear as a single curve for each value of ρ. (b) Comparing the $L(s)$ function for the DG and the $\tilde{L}(s)$ function for the EIF (after transforming from the EIF probability density function for the variable $\mathcal{S}$ to the DG variable s. The functions agree to an extent over the pdf $\phi_D G$ of the DG model. The DG function $L(s)$ tends to 0 for negative values of s where the LNL function $\tilde{L}(s)$ tends to a finite non-zero value. This agrees with the probability distributions in the previous model where the LNL cascade is less accurate at estimating the 0-population spike count and large population spike count probabilities. (c) The heat capacity increases linearly for the LNL-cascade, the EIF and the DG. For the case of the LNL cascade the heat capacity increases at a slightly greater rate than the EIF/DG which overlap. The Ising model saturates for a population of approximately $N = 30$ neurons.

the windowed firing rate $\mathcal{S}$ – each of the N neurons produces spikes independently. Thus, the probability of k cells firing simultaneously is

$$P_{\mathrm{LF}}(k) = \binom{n}{k} \int_{-\infty}^{\infty} \phi_{LF}(\mathcal{S}) \big(1 - \tilde{L}\big)^{n-k} \tilde{L}^k \mathrm{d}\mathcal{S} \qquad (2)$$

$\phi_{LF}(\mathcal{S})$ is the probability density function for $\mathcal{S}$. We estimate ϕ_{LF} numerically.

Figure 4(a) shows that the LNL cascade captures the general structure of the EIF population output across a range of population sizes. In particular, it produces an order-of-magnitude improvement over the PME model (see DJS values in Fig. 4(b)), and reproduces the skewed structure produced by beyond-pairwise correlations.

This said, the LNL model does not produce a perfect fit to the EIF outputs, the most obvious problem being the overestimation of the zero spike probabilities, which $N = 100$ case are overestimated by almost 100% (the tail probabilities are also underestimated). Notably, the LNL fits become almost perfect for lower correlations i.e. $\rho = 0.05$ (see supplemental material [E: Can we add such a panel for N=100 to supplemental ... like the N=100 histograms, but for rho=0.05?]). This suggests the errors are due to the way the static non-linearity deals with fluctuations to very low or high firing rates $r(t)$; these fluctuations are smaller at lower correlation values, which lead to smaller signal currents in the LNL formulation.

Relating the EIF and Dichotomous Gaussian models The LNL model provides a reduction of the EIF model to an inhomogeneous poisson process that is based directly on the underlying SDEs. In many settings – and especially in experiments – we seek more abstracted statistical models of spike outputs. We now demonstrate that one such approach, which has been shown to produce [? ?] and capture [?] higher-order correlations before, reproduces the EIF population activity with exquisite accuracy.

In this Dichotomous Gaussian framework, introduced by [? ?], N neurons receive correlated Gaussian input with mean γ and covariance λ. Each neuron applies a step nonlinearity to its inputs, spiking only if its input is positive. The mean of the input is chosen so that the mean of the output is μ and similarly λ is chosen so that the correlation coefficient is ρ. The correlated Gaussian can be written as: $Z_i = \gamma + \sqrt{1-\lambda} T_i + \sqrt{\lambda} S$ where T_i is the independent input and S is the common input. The probability of a spike is given by $P(Z_i > 0|s)$ and again we can define the $L(s)$ function: [E: check lower case vs upper case s usage]

$$L(s) = P\left(T_i > \frac{-s-\gamma}{\sqrt{1-\lambda}}\right) = \Phi\left(\frac{s+\gamma}{\sqrt{1-\lambda}}\right) \qquad (3)$$

[E: Φ used differently in diff parts of paper – use different choice for LNL model?] Equipped with Eqn. (3), the probability of observing a spike count k is the same as equation [eqnum] using $L(s)$ and $\phi_{DG}(s)$ is the probability density function of a one-dimensional Gaussian with mean 0 and variance λ.

Fig. 5a shows that the Dichotomized Gaussian model provides an essentially exact description of the EIF population output for a range of firing statistics. We evaluate this connection to the EIF model via the probability distributions P_{LNL} and P_{DG}. To make the comparison we must transform from the probability density function of the linear-nonlinear model ϕ_{LF} to the Gaussian pdf ϕ_{DG} using the nonlinear change of variable:

$$\mathcal{S} = f(s), \quad \text{where} \quad f'(s) = \frac{\phi_{DG}(s)}{\phi_{LF}\big(f(s)\big)}. \qquad (4)$$

Writing the LNL cascade probability in terms of the s variable we get:

$$P_{\mathrm{LF}}(k) = \binom{n}{k} \int_{-\infty}^{\infty} \phi_{\mathrm{DG}}(s)\big(1-\tilde{L}(f(s))\big)^{n-k}\tilde{L}(f(s))^{k}\mathrm{d}s \tag{5}$$

where $\tilde{L}(\mathcal{S}) = 1 - \exp(-\mathcal{S})$. After this transformation the only difference between the models is now the $L(s)$ functions. The comparison between L and $\tilde{L}$ can be seen in figure 4(b). The functions largely agree over about 2 standard deviations of the Gaussian pdf of values of the common input signal s. At large values of common input s, the higher values of the DG $L(s)$ account for the more accurate fit of the tail of $P(k)$ [E: Need to make sure phrasing and references to symbols here make sense.]

The success of the DG model in capturing EIF statistics is significant for two reasons. First, it suggests why this abstracted model has been able to capture the population output recorded from spiking neurons. Second, because the DG model is a special case of a Bernoulli generalized linear model (see supplemental material), our finding indicates that this very broad and easily fittable class of statistical models may be able to capture the higher-order population activity in neural data.

[E: As I am reading it at least I don't find the following convincing – 10^{-6} is seriously tiny! I say we just remove the following text.] For negative values of common input the DG $L(s)$ function approaches 0 whereas the LNL function reaches a small but finite value of approximately 10^{-6} (this is actually the initial condition in solving the ODE to find the nonlinear transformation $f(s)$ i.e. $f(0) = 10^{-6}$. This explains why the LNL cascade estimates the zero spike probabilities so poorly. Solving this ODE for f still needs to be checked!

[E: Add discussion of heat capacity ...]

Summary and conclusion: We have shown that Exponential-Integrate and Fire (EIF) neurons receiving common input give rise to strong higher-order correlations. Moreover, the correlation structure that results can be predicted from a linear-nonlinear cascade model, which forms a tractable reduction of the EIF neuron. Overall, the cascade model gives an order of magnitude better description of the EIF population spike count distribution and provides a simple mechanism for higher-order correlations. Moreover, this model can be directly related to the Dichotomized Gaussian model , which has been highly successful in the experimental and statistical literature.

The authors thank Liam Paninski for helpful insights. This work was supported by the Burroughs Wellcome Fund Scientific Interfaces Program and NSF Grants DMS-xxxx and CAREER-xxxxx.

[1] Shun-Ichi Amari, Hiroyuki Nakahara, Si Wu, and Yutaka Sakai. Synchronous firing and higher-order interactions in neuron pool. *Neural Comp.*, 15(1):127–42, 2003.

[2] Jakob H Macke, Manfred Opper, and Matthias Bethge. Common Input Explains Higher-Order Correlations and Entropy in a Simple Model of Neural Population Activity. *Physical Review Letters*, 106(20):208102, May 2011.

[3] Shan Yu, Hongdian Yang, Hiroyuki Nakahara, Gustavo S Santos, Danko Nikolić, and Dietmar Plenz. Higher-order interactions characterized in cortical activity. *Journal of Neuroscience*, 31(48):17514–17526, November 2011.

.

[E: Eventually break this off and submit as separate file ... typeset with single column]

SUPPLEMENTARY MATERIAL

Relating DG and Generalized Linear models: The LNL model provides a reduction of the EIF model to an inhomogeneous poisson process that is based directly on the underlying SDEs, and is in extremely wide use in neural modeling [**?**]. However, it far from the only approach to statistical modeling of spiking neurons. In particular, generalized linear models can be fit to the Bernoulli data given by the 1's and 0's of binned spikes in individual cells. Such models similarly apply a linear filter to the common input signal, and followed by a static nonlinearity $f(\cdot)$, to yield a spiking probability for the current time bin. Noting that any linear filter on our (gaussian white noise) input signal will yield a gaussian value s, this class of models therefore yields spiking probabilities $f(s)$ where s is gaussian. Comparing with Eqn. (3) in the main text, we see that the DG and GLM models have the same general form, when f is taken to be the cumulative distribution function for a gaussian (as in "probit" models).