HHS Public Access

Author manuscript

IEEE Trans Control Netw Syst. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:

IEEE Trans Control Netw Syst. 2018 September ; 5(3): 1146–1156. doi:10.1109/TCNS.2017.2687824.

Control Analysis and Design for Statistical Models of Spiking Networks

Anirban Nandi [Student Member, IEEE],
Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO-63130, USA.

MohammadMehdi Kafashan [Member, IEEE], and
Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO-63130, USA.

ShiNung Ching [Member, IEEE]
Department of Electrical and Systems Engineering, Biology and Biomedical Sciences, Washington University in St. Louis, MO-63130, St. Louis, shinung@ese.wustl.edu

Abstract

A popular approach to characterizing activity in neuronal networks is the use of statistical models that describe neurons in terms of their firing rates (i.e., the number of spikes produced per unit time). The output realization of a statistical model is, in essence, an n–dimensional binary time series, or pattern. While such models are commonly fit to data, they can also be postulated *de novo*, as a theoretical description of a given spiking network. More generally, they can model any network producing binary events as a function of time. In this paper, we rigorously develop a set of analyses that may be used to assay the controllability of a particular statistical spiking model, the point-process generalized linear model (PPGLM). Our analysis quantifies the ease or difficulty of inducing desired spiking patterns via an extrinsic input signal, thus providing a framework for basic network analysis, as well as for emerging applications such as neurostimulation design.

Keywords

Neural Control; PPGLM; Stimulation

I. Introduction

Understanding the control properties of brain networks at spatial scales commensurate with individual neurons has two important implications in neuroscience. First, the advent of technologies such as optogenetics [1] allow, in principle, for fine spatial actuation of networks at the scale of tens to hundreds of individual neurons. Achieving spatially and temporally precise control in such networks would provide a substantial tool in the probing of neural function [2]. Second, knowing the control properties of such networks may aid in understanding basic issues around how networks of neurons intrinsically self-coordinate their activity.

A direct approach to understanding the control properties of spiking neuronal networks (as distinct from macro-scale networks at the level of brain regions) involves the use of dynamical-systems models, such as the simple integrate-andfire neuron [3], or more detailed biophysical models involving voltage-gated conductance equations [4]. However, while basic control characterizations have been obtained in single neurons [5] and pairs of neurons [6], the nonlinearity and/or discontinuity associated with the neuronal dynamics in question lead to issues of scalability in both control analysis and design. It is important to make a distinction here between control at the level of asynchronous timed spiking, versus control at the level of population activity such as oscillations or synchronization, where control analysis results have been obtained [7], [8].

Statistical models provide a way to approximate underlying neural dynamics, wherein nonlinearity and inherent stochasticity are embedded within a time-varying rate function. The rate function associated with a neuron governs its probability of producing a spike at any moment in time, such as in the Poisson class of random processes. Thus, in a statistical model, neuronal spikes are described as binary events in a particular output realization. Such models include the popular class of point-process generalized linear models (PPGLMs) [9], [10], which have been used to model event-based phenomena in ecology [11], [12], telecommunications [13] and, in the present context, the spiking activity in neuronal networks. Here, each spike is understood as a timed binary event. Since they are readily fit to spiking data, PPGLMs have emerged as a powerful tool in the analysis of neural recordings [14]. Absent data, PPGLMs can also be formulated *de novo* as mathematical models of neural activity that can capture some aspects of the network structure and dynamics (e.g., delays, refractory).

Recently, the problem of extrinsic neural control has been formulated for this class of models by including an extrinsic input as a covariate in the PPGLM [15]. However, as yet, no methods for basic control analysis have been developed for such models. Such analysis is needed in order to provide baseline characterizations such as establishing whether or not a design objective is feasible. For instance, it would be vacuous to attempt a design on a system that was not theoretically controllable. The goal of this paper is to bridge this gap by providing a set of quantitative metrics, based on dynamic optimization, that assay the control properties of a statistical neural model, to enable basic characterizations (e.g., given two GLMs, which is ‘more controllable’) and, eventually, the problem of input design. In this vein, we provide four main results: (i) a control analysis for PPGLMs that approximates, in essence, the reachable set of binary patterns for a given model; (ii) from (i), a relativistic notion of control viability that allows comparison between PPGLMs; (iii) a validation of the proposed framework, showing its ability to reveal salient control properties of spiking networks; and (iv) the instantiation of the developed theory for the purposes of designing external neurostimulation. It is important to note that our control analysis is developed strictly for PPGLMs, which, as stated above, are approximations of the true dynamics.

The remainder of the paper is organized as follows. In Section II we discuss the PPGLM models used in this paper. Next, in Section III, we analyze the likelihood function with respect to the extrinsic control and show the role of event count in determining the maximum achievable likelihood. We introduce a controllability-like notion based on this

count and the likelihood that estimates the pattern subset that can be viably obtained in terms of probability from the set of all possible patterns. Section IV provides numerical simulations to validate our analysis and demonstrates how the analysis can be used to compare two different point process models. Finally, in Section V, we use the analysis results to construct a control design framework to induce any desired spike pattern.

II. Preliminaries

In this section we first demonstrate the intuition behind using a PPGLM to model the spiking activity in neurons and then proceed to develop the probabilistic descriptions of patterns of activity in these neurons.

A. Notation

A point process is an integer-valued stochastic process that models the occurrence of isolated events in time and space, e.g., neural spiking. The inhomogeneous Poisson process is one such example that can capture temporal dependencies via a time-varying rate/intensity function [16]. Generalized Linear Models (GLMs) provide a regression framework to model output variables Y with respect to the input/explanatory variables X. GLMs assume that a transformation of the conditional mean of Y is a linear function of X, i.e.,

$$g(\mathbb{E}(Y|X)) = X\beta \quad (1)$$

where $g(.)$ is the link function and β is a set of unknown parameters. Combining point processes with GLMs, i.e., modeling the rate function of neurons by a GLM, results in a PPGLM, the primary object of study in this paper.

Throughout this paper, events and spikes are used synonymously. Most mathematical notation is standard. The continuous time univariate and multivariate point processes are indicated by $N(t)$ and $\mathbf{N}(t)$, $t \in \mathbb{R}^+$ respectively, whereas $N_{t'}$ and $\mathbf{N}_{t'}$, $t' \in \mathbb{N}$ denote their discrete counterpart. In a univariate discrete process, N_i, the value of discrete process at the i-th window, is a scalar. For a multivariate discrete process, $\mathbf{N}_i$ is a vector of all variables at the i-th window and $N_{c,i}$ is a scalar that represents the value of the c-th variable in the i-th time bin. We follow the same notation for the associated difference processes.

B. Model Description (Exclusive Event Point Process)

We first consider a univariate inhomogeneous Poisson process $N(t)$ with the intensity (event rate) function $\lambda(t|H(t))$, where $H(t)$ denotes the history of the process along with other covariates, i.e.,

$$\lambda(t|H(t)) = \lim_{\Delta \to 0} \frac{\Pr[N(t+\Delta) - N(t) = 1|H(t)]}{\Delta}. \quad (2)$$

We divide the total time window under consideration, $[0, T]$, into I intervals such that $\Delta = T/I$ and denote the discrete process as $N_i \equiv N(i\Delta)$ and $H_i \equiv H(i\Delta)$, $i = 1,...,I$. This yields the difference process

$$\delta N_i = N_i - N_{i-1} = N(i\Delta) - N((i-1)\Delta). \quad (3)$$

We make the key assumption that $\Delta \ll 1 (\Delta \neq 0)$, resulting in $\delta N_i \in \mathbb{B}$, where $\mathbb{B} := \{0,1\}$. We separate the conditional intensity (2) into components related to the background activity, spiking history over Q lags, and S independent extrinsic control inputs $\mathbf{U} \in \mathbb{R}^{S \times I}$, up to P previous instances via the log-link model

$$\lambda_i \equiv \lambda(i\Delta \mid \mathbf{X}, H_i) = \exp(\beta_0 + \sum_{q=1}^{Q} \beta_q \delta N_{i-q} \quad (4)$$
$$+ \sum_{s=1}^{S} \sum_{p=0}^{P} \gamma_p^s u_{s,i-p}) = \exp(\theta^T \mathbf{x}_i).$$

The parameter set is given by $\theta = \left[\beta_0 \ldots \gamma_P^S\right]^T \in \mathbb{R}^F, F = 1 + Q + (P+1)S$ and the co-variate matrix $\mathbf{X} \in \mathbb{R}^{F \times I}$ with the i-th column $\mathbf{x}_i$ as

$$\mathbf{x}_i = [1\ \delta N_{i-1} \ldots \delta N_{i-Q} u_{1,i} \ldots u_{S,i-P}]^T, \quad (5)$$

$\forall i = 1...I.$

The joint likelihood of a particular realization of $N(t)$ with k spikes over the I intervals, conditioned on $\mathbf{X}$, follows the form detailed in [17]

$$\Pr(\mathbf{N} \mid \mathbf{X}) = \exp\left(\sum_{c=1}^{C} \sum_{i=1}^{I} \delta N_{c,i} \log\left(\lambda_{c,i}\Delta\right) - \lambda_{c,i}\Delta \right) + o\left(\Delta^k\right). \quad (6)$$

where any function $f(x) \in o(h(x))$ implies that $\lim_{x \to 0} \frac{f(x)}{h(x)} \to 0$. We extend this model for the C-variate process $\mathbf{N}(t)$ as in [18] and the log-likelihood for small Δ can be written as

$$L(\mathbf{N} \mid \mathbf{X}) \equiv \log(\Pr(\mathbf{N} \mid \mathbf{X})) \quad (7)$$
$$= \sum_{c=1}^{C} \sum_{i=1}^{I} \left(\delta N_{c,i} \log\left(\lambda_{c,i}\Delta\right) - \lambda_{c,i}\Delta\right),$$

where

$$\lambda_{c,i} = \exp(\beta_0^c + \sum_{c'=1}^{C} \sum_{q=1}^{Q} \beta_q^{c',c} \delta N_{i-q} + \sum_{s=1}^{S} \sum_{p=0}^{P} \gamma_p^s u_{s,i-p}).$$

In terms of neural spiking, this set of co-variates captures:

1) any baseline activity in the network, via the bias term $\left(\beta_0^c\right)_{c=1}^{C}$;

2) refractory periods following a spike in the c-th neuron, via the self process history $\left(\beta_q^{c,c}\right)_{q=1}^{Q}$;

3) afferent excitation or inhibition from other neurons, via the network spiking history $\left(\beta_q^{c,c'}\right)_{c'=1,c'\neq c}^{C}$;

4) temporal dynamics (e.g., exponentially decaying) of the excitation or inhibition from other neurons, via additional history terms $\left(\beta_q^{c,c'}\right)_{c'=1,q=1}^{C,Q}$; and

5) effect of any extrinsic stimulation and the integrative nature in which the neurons process such information, via the current stimulus and the history terms $\left(\gamma_p^{p,s}\right)_{p=0,s=1}^{P,S}$.

However, more detailed biophysical dynamics associated with sub-threshold membrane potential and particular ion channels are outside the explanatory power of this model.

C. Model Description (Simultaneous Event Point Process)

The model described in (2)–(7), albeit useful in many contexts, is limited because it excludes multiple neurons producing simultaneous spiking events. Thus, we also consider a second model, a discrete-time, multinomial generalized linear model of a simultaneous event multivariate point process (SEMPP) [19], [20]. The coincidence of spiking events (simultaneous events) from different neurons in the interval Δ, is handled by projecting the system onto higher dimensions such that only a single kind of event can occur at any interval.

Briefly, for a C-dimensional inhomogeneous Poisson process $\mathbf{N}(t)$, a new $M = 2^C - 1$ dimensional marked point process $\mathbf{N}^*(t)$ is defined such that at any interval, there is at most one non-zero bit. The conditional intensity function for this marked point process $\mathbf{N}^*(t)$ is defined as $\lambda_m^*(t|H(t))$, $m = 1,...,M$, similar to (2) where $H(t)$ denotes the history of the process along with other covariates.

Once again with $\Delta = T/I \ll 1$ ($\Delta \neq 0$) over the time window $[0, T]$, we denote the discrete process as $N_{c,i}$ for $c = 1...C$ which yields the difference process $\delta N_{c,i}$ (for the multivariate point process), $\delta N_{m,i}^* \in \mathbb{B}$ (for the marked point process) similar to (3). In matrix representation we can write

$$\delta \mathbf{N} = \mathbf{D}\mathbf{N}, \quad (8)$$

where $\mathbf{D} \in \mathbb{R}^{I \times I}$ transforms $\mathbf{N}$ to its difference process $\delta\mathbf{N}$ (similarly for the marked process $\mathbf{N}^*$). Here a logistic-link function is used to relate the co-variates with the rate of the process,

$$\begin{aligned} \log \frac{\lambda^*_{m,i}\Delta}{1 - \lambda^g_i \Delta} &= \beta^m_0 + \sum_{c=1}^{C} \sum_{q=1}^{Q} \beta^{m,c}_q \delta N_{c,i-q} \quad (9) \\ &+ \sum_{s=1}^{S} \sum_{p=0}^{P} \gamma^{m,s}_p u_{s,i-p} \\ &= \theta^T_m \mathbf{x}_i, \end{aligned}$$

where $\lambda^g_i = \sum_{m=1}^{M} \lambda^*_{m,i}$ is the conditional intensity for the discrete ground process [21] $N^g_{t'}$ at $t' = i.\theta_m$ is the m-th row of the parameter matrix $\Theta \in \mathbb{R}^{M \times F}$ with

$$F = 1 + QC + (P+1)S \quad (10)$$

co-variates at each interval. Θ reflects the dependence of the intensity function on the co-variates $\mathbf{X} \in \mathbb{R}^{F \times I}$. The loglikelihood for the marked point process conditioned on the co-variates $\mathbf{X}$, is given by

$$\begin{aligned} L(\mathbf{N}^* | \mathbf{X}) = \log\left(\Pr(\mathbf{N}^* | \mathbf{X})\right) &= \sum_{i=1}^{I} \sum_{m=1}^{M} \delta N^*_{m,i} \theta^T_m \mathbf{x}_i \quad (11) \\ - \sum_{i=1}^{I} \log\left(1 + \sum_{m=1}^{M} \exp\left(\theta^T_m \mathbf{x}_i\right)\right). \end{aligned}$$

In the analysis that follows, we work with both the likelihood models in (7), (11). Much of the analysis that follows will be based on characterizing how the number of events (spikes) in a target realization impacts these likelihoods. We specifically consider the spike count $\psi : \mathbb{R}^{C \times I} \rightarrow \mathbb{R}$,

$$\psi(\delta\mathbf{N}) = \psi(\mathbf{N}\mathbf{D}) = \mathbf{b}^T \delta \mathbf{N} 1_I = \sum_{c=1}^{C} \sum_{i=1}^{I} b_c \delta N_{c,i}, \quad (12)$$

as the number of events in the realization. For the exclusive event process $\mathbf{b} = \mathbf{1} \in \mathbb{R}^C$ and (12) reduces to

$$\psi(\delta\mathbf{N}) = \sum_{c=1}^{C}\sum_{i=1}^{I}\delta N_{c,i}. \quad (13)$$

For SEMPP, $\mathbf{b} \in \mathbb{R}^M$ contains the number of events associated with each dimension of the projected point process, e.g., for $C = 3$, the projected dimension is $M = 7$ and

$$\mathbf{b} = [1\ 1\ 2\ 1\ 2\ 2\ 3]^T, \quad (14)$$

corresponding to all possible combinations, i.e., three 1-spike events, three 2-spike events and one 3-spike event.

III. Control Analysis of Statistical Spiking Models

In this section based on the likelihood models developed above we approach the question of controllability in spiking networks from a probabilistic standpoint. In particular we identify spike count as a key marker that relates to the probability of achieving any spike pattern as a function of extrinsic control.

A. ϵ -Controllability for PPGLMs

We first consider an analogue to the classical notion of controllability. As a statistical model, any such notion must involve the likelihood of particular realizations, heretofore referred to as *target patterns*. As such, we first consider the following candidate:

Definition 1 (ϵ–Controllability for PPGLMs): A PPGLM is −controllable if, for all > 0, there exists an input U such that any realization $\mathbf{N}(t)$ of the PPGLM can attain a loglikelihood satisfying

$$-\epsilon \leq L(\mathbf{N}|\mathbf{U}) \leq 0. \quad (15)$$

Despite its intuitive appeal, the following highlights that the notion of −controllability is too strong to be of practical utility in the desired context.

Lemma 1: The PPGLM described in (7), (11) is not **ϵ**− controllable, even if the energy of the input **U** is unconstrained.

Proof 1: The proof is given in Appendix A, and hinges on the fact that the likelihood function is in fact strictly concave in **U**.

The Lemma establishes that allowing **U** to assume arbitrarily large energy confers no advantage in controlling the PPGLM. This is conceptually different from classical control analysis, where allowing progressively larger energy (in general) improves the overall range of trajectories that can be induced. Two points should be considered when interpreting this

result. First, our analysis focuses on at most one event in each time bin. With increasing energy, one may increase the likelihood on *an* event, but not necessary a single one. Second, in a coupled network scenario, applying a large input in order to target a spike in a particular neuron will have collateral effects elsewhere in the network. However, clearly some *minimum* energy is required in order to maximize the likelihood of given realizations.

B. Event Count as a Surrogate for Pattern Complexity

As a consequence of Lemma 1, we seek a characterization that examines the complexity of the realizations (spike patterns) that can be induced. Below, we establish that the spike count, i.e. simply the number of spikes contained in a particular realization (i.e. (12)), can serve as an informative marker in this regard.

Lemma 2: For a PPGLM of the form (4)–(7) with infinitesimally small interval ($\Delta \ll 1$), the maximum likelihood of any realization decreases with respect to number of events.

Proof 2: The proof is given in Appendix B.

Lemma 2 is most easily understood in a *fully actuated* scenario wherein each neuron receives its own, independent control input. In this case, it is straightforward to show that the control can be designed to negate any effect of process history. Consider the likelihood model of (7) with $S = C$ along with $P = 0$ and $\gamma_0^{c,s} = 0$ for $c \neq s$, i.e., Γ_0, which reflects how the current input affects all the processes, is a C-dimensional vector. Since here the probability of an event is independent at each time and other input indices, we can analyze the likelihood for each i and c separately, i.e.,

$$\max_{U \in \mathbb{R}^{C \times I}} L(\mathbf{N}|\mathbf{U}) = \sum_{i=1}^{I} \sum_{c=1}^{C} \max_{u_{c,i} \in \mathbb{R}} L\left(N_{c,i}|u_{c,i}\right) \quad (16)$$
$$= \sum_{i=1}^{I} \sum_{c=1}^{C} \max_{u_{c,i} \in \mathbb{R}} L_{c,i},$$

where

$$L_{c,i} \equiv L\left(N_{c,i}|\mathbf{u}_i\right) = \delta N_{c,i}\left(\theta_c^T \mathbf{x}_i + \log \Delta\right) - \Delta \exp\left(\theta_c^T \mathbf{x}_i\right). \quad (17)$$

For $\delta N_{c,i} = 0$, (17) reduces to

$$L_{c,i} = -\Delta \exp\left(\theta_c^T \mathbf{x}_i\right) = -\Delta \exp\left(r_{c,i} + \gamma_0^c u_{c,i}\right). \quad (18)$$

We observe that given any $\epsilon > 0$,

$$L_{c,i}\left(\delta N_{c,i}=0\,|\,u\right)\geq\ -\epsilon'\ \text{ when }\ u\leq u^*_{c,i},\quad (19)$$

where $u^*_{c,i}=\frac{1}{\gamma^c_0}\left(\log\left(\frac{\epsilon'}{\Delta}\right)-r_{c,i}\right)$, assuming $\gamma^c_0>0$. In other words, for the pattern consisting of all zeros, the likelihood indeed can be made arbitrarily close to one (for the fully actuated case). We now show that the addition of any spike to the pattern results in likelihood degradation.

Specifically, for $\delta N_{c,i}=1$, we can maximize the indexed likelihood as

$$L_{c,i}\left(\delta N_{c,i}=1\,|\,u^*_{c,i}\right)=\ -1\ \text{ with }\ u^*_{c,i}=\ -\frac{1}{\gamma^c_0}\left(\log\Delta+r_{c,i}\right).\quad (20)$$

Since the maximum likelihood of each $L_{c,i}$ is fully determined by the input $u_{c,i}$, we can design an extrinsic control $\mathbf{u}^*$ using (19), (20) that maximizes the likelihood for the whole realization, i.e. from (16)

$$\begin{aligned}L(\mathbf{N}\,|\,\mathbf{U}^*)=&\sum_{\delta N_{c,i}=0}L_{c,i}\left(\delta N_{c,i}=0\,|\,\mathbf{U}^*\right)+\quad (21)\\&\sum_{\delta N_{c,i}=1}L_{c,i}\left(\delta N_{c,i}=1\,|\,\mathbf{U}^*\right)\\=\ &-(CI-\psi(\delta\mathbf{N}))\epsilon'-\psi(\delta\mathbf{N})\end{aligned}$$

Now for unconstrained inputs we have

$$\lim_{\epsilon'\to 0}L(\mathbf{N}\,|\,\mathbf{U}^*)\approx\ -\psi(\delta\mathbf{N}),\quad (22)$$

i.e., the maximum likelihood decreases with the number of events $\psi(\delta\mathbf{N})$ in any realization $\mathbf{N}$ of the process $\mathbf{N}(t)$.

A similar analysis can be carried out for the fully actuated SEMPP model, wherein we can treat marked process independently and use (16) to maximize the likelihood over the whole realization. In this case, the likelihood at the c-th process, i-th time index is

$$L_{c,i}\equiv L\left(N_{c,i}\,|\,\mathbf{u}_i\right)=\delta N_{c,i}\gamma u_{c,i}-\log\left(1+\exp\left(\gamma' u_{c,i}\right)\right).\quad (23)$$

Since the inputs are unconstrained, the total contribution from the co-variates can be reformulated in terms of only two parameters $\boldsymbol{\gamma}$, $\boldsymbol{\gamma}'\ \forall\ c, i$. Also note that we have removed the asterisk indicating the marked point process since analyzing each process independently

in one dimension means, $\delta N^*_{c,i} = \delta N_{c,i}$. For $\delta N_{C,i} = 0$, we can achieve probability approximately close to one and can obtain a similar version of (19). When $\delta N_{C,i} = 1$, the maximum is attained at

$$L_{c,i}\left(\delta N_{c,i} = 1 \mid u^*_{c,i}\right) = \varphi(\gamma, \gamma') \text{ with } u^*_{c,i} = \frac{1}{\gamma'} \log\left(\frac{\gamma}{\gamma' - \gamma}\right), \quad (24)$$

where

$$\varphi(\gamma, \gamma') = \log\left(\frac{\gamma^{\frac{\gamma}{\gamma'}}}{\gamma'(\gamma' - \gamma)^{\frac{\gamma}{\gamma'} - 1}}\right) < 0. \quad (25)$$

The likelihood maximization in (24) is independent of each c, i and similar to (21), we have

$$L(\mathbf{N} \mid \mathbf{u}^*) \approx \psi(\delta \mathbf{N}) \varphi(\gamma, \gamma'). \quad (26)$$

Thus, from our analysis of both the likelihood models (7), (11), we can conclude that in terms of likelihood, increasing the number of spikes in a pattern results in likelihood degradation, which can be interpreted as greater control difficulty.

C. Estimation of Complexity-based Viable Sets

Clearly, there are many factors in addition to spike count that determine the likelihood of a particular realization of a considered PPGLM. Indeed, not all patterns with the same spike count will generate the same likelihood. Accepting this limitation (see also **Section VI**), we will leverage the result of the previous section to form a tractable, accurate assay for the control properties of a PPGLM in terms of spike count. We proceed first by introducing the notion of a viable pattern set, which is analogous to the reachable set for a classical control system.

Definition 2 (ρ-Viable Pattern Set): Consider an arbitrary M-dimensional PPGLM defined over I intervals. Given a likelihood threshold ρ, the ρ-Viable Pattern Set, $\mathcal{N}(\rho; C, I, \mathcal{U})$, is the set of patterns defined as

$$\mathcal{N}(\rho; C, I, \mathcal{U}) = \left\{\mathbf{N} \in \mathbb{R}^{C \times I} \mid \exists \mathbf{U} \in \mathcal{U} \text{ s.t. } \Pr(\mathbf{N} \mid \mathbf{U}) \geq \rho\right\}, \quad (27)$$

where $\mathcal{U}$ denotes the set of admissible inputs.

It follows from Lemma 2 that, in general, $\mathcal{N}(\rho; C, I, \mathcal{U})$ includes all patterns with a spike count less than or equal to some maximally viable count, $\mu \leq CI$. We can thus formulate a relativistic analysis as follows.

Definition 3 ((μ,ρ)-Viability): For a likelihood threshold ρ and spike count μ, the PPGLM (4)–(7), is (μ,ρ)-viable if $\exists\ u \in \mathcal{U}$ such that

$$\mathcal{N}(\rho; C, I, \mathcal{U}) \supset \mathcal{N}_{\mu}(C, I), \quad (28)$$

where $\mathcal{N}_{\mu}(C, I)$ denotes the set of all patterns with spike counts of μ or less, i.e., $\forall\ \mathbf{N} \in \mathcal{N}_{\mu}(C, I)$, we have

$$\psi(\delta\mathbf{N}) = \mathbf{b}^T \delta\mathbf{N}\ \mathbf{1} \leq \mu \quad (29)$$

where $\delta\mathbf{N} \in \mathbb{B}^{C\times I}$ is the difference process corresponding to $\mathbf{N}$.

The key problem is now to obtain the maximally viable count, μ, for a given ρ. This amounts to a joint optimization problem for the spike count, $\psi(\delta\mathbf{N})$, and control $\mathbf{U}$. Since the difference process imposes the constraint

$$\delta N_{c,i} = \{0, 1\}, \quad \forall c, i, \quad (30)$$

this optimization is a Nonlinear Mixed Integer program. To make this tractable, we relax the integer constraint and introduce a new variable χ, such that

$$\chi_{c,i} \in [0, 1], \quad \forall c, i\,. \quad (31)$$

This allows us to define a relaxed viability notion as follows.

Definition 4 (Relaxed maximally viable spike count): The relaxed maximally viable spike count μ_r is defined as

$$\mu_r = \psi(\chi) \quad (32)$$

and can be calculated from the solution of the following program,

$$\begin{aligned} \underset{\mathbf{U},\chi}{\text{maximize}}\ & \psi(\chi) \\ \text{subject to}\ & L(\chi|\mathbf{U}) \geq \log(\rho) \\ & u_{s,i} \in \mathcal{U},\ \forall s, i \\ & 0 \leq \chi_{c,i} \leq 1 \quad \forall c, i\,. \end{aligned} \quad (33)$$

While this optimization is still non-convex, we show below that numerical evaluation of the pairs (μ_r,ρ) leads to accurate, informative characterization of PPGLMs.

IV. Validation of the Analysis Framework

Here through numerical simulations we verify how relaxed maximally viable spike count affects the probability of achieving any pattern.

A. (μ_r,ρ)-Viability is Accurate

Figure 1 demonstrates the veracity of the relaxation in (33). We consider PPGLMs with randomly selected parameters Θ for $C = 3$ neurons, $Q = 6$ lags and $I = 10$ time bins. The inputs are constrained via $\mathcal{U} = [-5, 5]$. We solve (33) numerically[1] for the likelihood in (7) and find the relaxed maximum spike count $\mu_r = 7.08$ for $\rho = 10^{-6}$, $\Delta = 0.1$ and one independent input, i.e., $S = 1$. Then, the maximum likelihood is calculated individually for 100 random patterns and compared to the results of the (μ_r,ρ) optimization. Only 8⁄100 patterns are misclassified (spike counts that are below μ_r but nevertheless whose likelihoods do not exceed ρ). Patterns whose spike counts exceed μ_r are always classified correctly in this example.

B. (μ_r,ρ)-Viability Enables Salient Comparison of PPGLMs

Based on our formulation of (μ,ρ)-viability, if $\mu_r^1 > \mu_r^2$,

$$|\mathcal{N}^1(\rho)| > |\mathcal{N}^2(\rho)| \quad (34)$$

where |.| denotes the cardinality of a set. We demonstrate the utility of the viability analysis via an example, where we show how the analysis can disassociate PPGLMs with symmetric and asymmetric connectivity (see Figure 2A). We consider PPGLMs with the same structure, input constraint and window length as in the previous example and a fixed reference parametrization (essentially, the connectivity between units) $\Theta = \sigma \times \Theta_r$, where Θ_r is the base parameter. Two observations are of note in Figure 2B. First, a small amount of connectivity (via the scaling parameter $\sigma \in [0,1]$) is advantageous for control, beyond which viability decreases monotonically. This numerical inference can, in fact, be substantiated via a formal analysis:

Lemma 3: For a PPGLM modeling Exclusive or Simultaneous Event Processes with likelihoods defined in (7), (11) and connectivity defined via the parameters $\beta_q^{m,c}$, $\forall\ q,m,c$ ($M = C$ for the log-link model), the likelihood of any given pattern is strictly concave with respect to the network connectivity parameters.

Proof 3: The proof is contained in Appendix C, and is a variation of the proof of Lemma 1.

The second observation is that an asymmetric topology is, in general, more viable than a symmetric topology, consistent with studies of similar 3-neuron motifs using dynamical systems models and Lie bracket-based controllability analysis [22].

[1] We used a random sampling procedure over the initial conditions of our solver to ensure convergence to a robust local maximizer

V. Control Design of Statistical Spiking Models

The previous section focused on the development of the analytical framework for PPGLMs based on optimization. Along with this analysis, it is natural to also consider the overt design of an exogenous control input $\mathbf{U}^*$ to induce a specific target spiking pattern $\mathbf{N}_T$ with highest probability.

A. Control Design with Maximum Likelihood Estimation

Considering a cost function $\mathcal{J} : \mathbb{R}^{C\times I} \times \mathbb{R}^{C\times I} \to \mathbb{R}$, that accepts two patterns and returns a real number denoting how dissimilar they are, we can formulate the following optimization problem,

$$\begin{aligned}\mathbf{U}^* &= \arg\min_{\mathbf{U}\in\mathcal{U}} \left\langle \mathcal{J}(\mathbf{N},\mathbf{N}_T)\right\rangle_{\Pr(\mathbf{N}|\mathbf{U})} \\ &= \arg\min_{\mathbf{U}\in\mathcal{U}} \sum \mathcal{J}(\mathbf{N},\mathbf{N}_T)\Pr(\mathbf{N}|\mathbf{U}),\end{aligned} \tag{35}$$

where the sum is over all possible spike patterns $\mathbf{N}$. For a delta cost function $\mathcal{J}(\mathbf{N},\mathbf{N}_T) = -\delta(\mathbf{N},\mathbf{N}_T)$ as proposed in [15] we can rewrite (35) as

$$\begin{aligned}\mathbf{U}^* &= \arg\min_{\mathbf{U}\in\mathcal{U}} -\Pr(\mathbf{N}_T|\mathbf{U}) = \arg\min_{\mathbf{U}\in\mathcal{U}} -\log\ \Pr(\mathbf{N}_T|\mathbf{U}) \\ &= \arg\min_{\mathbf{U}\in\mathcal{U}} -L(\mathbf{N}_T|\mathbf{U}).\end{aligned} \tag{36}$$

Thus, the delta cost function reduces (35) to a maximum likelihood estimation (MLE) problem.

Proposition 1: The maximum likelihood estimation problem of finding the extrinsic control $\mathbf{U}^*$ for the likelihood defined in (7) and (11), under an energy constraint on the control, is convex.

Proof 4: In Lemma 1, we have established that the likelihoods presented in (7) and (11) are strictly concave with respect to the extrinsic control $\mathbf{U}$, which makes (36) convex.

Note that any constraint on energy (quadratic form) will not alter the convexity of the program. Also, any regularization in the cost in terms of energy effectively makes the problem a maximum *a posteriori* (MAP) estimation problem.

B. Analysis and MLE Design Example

1) Verification of the Controllability Analysis: Here we validate our controllability analysis results on a randomly parametrized PPGLM model equipped with a log-link function (4). First we solve for maximally viable spike count μ_r, (33) with $C = S = 4$, $\rho = 10^{-8}$, $\Delta = 0.01$ and $\mathcal{U} = [-10, 10]$. Then for some randomly chosen (μ_r,ρ)-viable pattern $\mathbf{N}_T$ (using (28) and (29)) shown in Figure 3 (top panel), we calculate $\mathbf{U}^*$ from the MLE problem

in (36). In the middle and bottom panel of Figure 3 we plot the probability of spike in each window ($\lambda_{c,i}\Delta$) and the corresponding extrinsic control input $u_{s,i}\ \forall\ c,s,i$ from the maximization solution. We also observe that indeed the pattern $\mathbf{N}_T$ is ρ-viable (27). We note that in this example, the low probability of spiking in Neuron 1 is due to the presence of large excitatory connectivity between Neuron 1 and 2. Thus, the MLE solution biases the resultant pattern in order to avoid spurious spiking in Neuron 2.

2) PPGLM Control of Underlying Stochastic Integrate and Fire (INF) model: Finally, in this section we illustrate that our design strategy can be used indirectly to control dynamical systems models. Here we consider C coupled stochastic integrate and fire (INF) neurons of the form [23],

$$\frac{dV(t)}{dt} = -\frac{1}{\tau_v}V(t) + \frac{1}{\mathscr{C}}\left(bu(t) + I_{syn}(t)\right) + \eta\, e(t) \quad (37)$$

$$I_{syn}(t) = -g_{syn}(t)\left(V(t) - E_{syn}\right)$$

$$g_{syn}(t) = \bar{g}\frac{(t - t_s)}{\tau_s}\exp\left(-\frac{(t - t_s)}{\tau_s}\right)$$

where τ_V is the membrane time constant, $\mathscr{C}$ is the membrane capacitance, $e(t)$ is standard Gaussian white noise, η denotes the standard deviation of this noise, $u(t) \in \mathbb{R}^S$ is the extrinsic control input, $b \in \mathbb{R}^{1\times S}$ denotes the influence of the input on the neuron, $I_{syn}(t)$ is the synaptic current coming from a pre-synaptic neuron firing an action potential at time t_s, E_{syn} is the reversal potential of the synapse, $\bar{g}$ models the constant synaptic conductance and τ_s determines the decay of the synaptic current as time is elapsed from the incoming spike at t_s. The model parameters for the neurons are given by

$$\mathscr{C} = 10\ nF,\ \tau_v = 15\ ms,\ V_{rest} = -70\ mV, \quad (38)$$
$$V_{Thresh} = -50\ mV,\ E_{syn} = 70\ mV,\ \eta^2 = 2$$
$$\bar{g} \sim \mathscr{U}[0,1],\ \tau_s = 1\ ms,\ b \sim \mathbb{N}(0,1)\,.$$

Now, we determine the GLM model parameters Θ in a Monte Carlo fashion in $K_T = 500$ different trials. Exciting stochastic INF network with $\mathbf{U}(t)$ drawn from a Gaussian distribution such that $u_{s,i}(t) \sim \mathbb{N}(0,50)\ \forall\ s,i$, produces spike patterns $\mathbf{N}_j$ for $j = 1...K_T$ and using these data we fit $\widehat{\Theta}$ that best describes the training set. Conceptually, this is akin to a system identification step.

In Figure 5 we show the performance of the control $\mathbf{U}^*$, obtained from the delta objective, on the INF network for different cases of actuation (C neurons, S inputs). The covariate matrix $\mathbf{X}$ has three process lags ($Q = 3$) and input history ($P = 5$), selected based on the Akaike information criterion (AIC). Figure 4 shows the Kolmogorov-Smirnov (KS) goodness-of-fit test using time-rescaling theorem [24], which indicates that the model accurately reflects the data. With the hypothesis that the control inputs calculated from the

PPGLM should also emit a spike train close to the target $\mathbf{N}_T$ (panel A) in the underlying dynamical model, we stimulate the INF neurons with the same **U*** (panel D). In panel B, C we show the generated spike pattern (averaged over several realizations) and one sample waveform, when the original stochastic INF neurons are excited by **U***, and indeed we can see that the induced pattern is close to target $\mathbf{N}_T$. For validation of optimality of **U***, in panel E we plot the achieved spike pattern for a randomly selected input. We observe that as underactuation becomes more prominent, the performance of **U*** degrades. The simulation results were primarily generated using CVX with MATLAB interface.

3) Control Design for underlying Biophysical Models: To evaluate the utility of this design approach on a more complicated biophysical model, we further consider a network of diffusively coupled Fitzhugh-Nagumo (FN) neurons of the form [25]. Here the dynamics of the c-th neuron is given by

$$\frac{dv_c}{dt} = v_c - \frac{v_c^3}{3} - w_c + bu(t) + \frac{\sigma_w}{C}\sum_{c'=1}^{C}\left(v_c - v_{c'}\right) + \eta\, e(t) \qquad (39)$$
$$\tau\frac{dw_c}{dt} = v_c + \bar{\alpha} - \bar{\beta} w_c,$$

where v_c denotes the membrane potential, w_c the recovery variable, σ_w is the coupling strength, α, β, τ are system parameters and $u(t)$, $e(t)$ the extrinsic input and standard Gaussian white noise respectively as before in (37). In our simulations we have used

$$\tau = 12.5,\ \bar{\alpha} = .7,\ \bar{\beta} = 0.8,\ \eta^2 = 0.5 \qquad (40)$$
$$\sigma_w \sim \mathcal{U}[0,1],\ b \sim \mathbb{N}(0,1)\,.$$

We use a spike detection algorithm that records a spike from simulated voltage for amplitudes higher than $V_{Thresh} \sim 1\ mV$ [26] and refractory period of 2 ms. In Figure 6 we show the average achieved pattern and one voltage waveform (panel B and C respectively) for a randomly selected target pattern (panel A) for a fully actuated ($C = S = 2$) network of FN neurons as in (39).

VI. Conclusions

In this paper, we have introduced a control analysis paradigm for statistical models. We show that the number of events in a realization of any PPGLM is a simple indicator of pattern viability in terms of likelihood. This enables us to formulate a problem of finding the subset of all possible patterns that can be realized with a specified probability – the viable pattern set. In turn, we can directly compare two PPGLMs with respect to this likelihood viability. Such an analysis provides an important means to compare the extent to which different PPGLMs can be controlled, and we demonstrated the accuracy of the proposed analysis. Finally, we show how the analysis can naturally pair with a design paradigm to compute optimal controls for inducing a desired pattern on the PPGLM.

It is important to note the limitations in the proposed approach. Most notably, we focus here on evaluating the (relaxed) maximally viable spike count μ_r to investigate the space of patterns that can be achieved to within the probability threshold ρ. As mentioned, this framework does not distinguish between different patterns with the same count μ, and labels all of them to be viable for ρ if $\mu \leq \mu_r$. For an idealistic scenario, i.e. full actuation and unconstrained control inputs, we proved in (21), (26) that the event count solely dictates likelihood degradation, but with stringent energy constraints on the input and heavy underactuation, the process history and other co-variates also affect the likelihood and so the dependence on $\psi(\delta\mathbf{N})$ is not exclusive. The misclassified patterns in Figure 1 are attributed to this fact.

Understanding this limitation, the aforementioned issue of non-convexity of (33) and the constraint relaxation (31), we posit that the framework is strong enough to reveal salient control properties in spiking networks. Our example highlighting concave dependence on connection strength, a fact that is analytically verifiable, demonstrates this utility.

Acknowledgment

S.C. holds a Career Award at the Scientific Interface from the Burroughs-Wellcome Fund. This work was supported by AFOSR 15RT0189, NSF ECCS 1509342 and NSF CMMI 1537015, from the US Air Force Office of Scientific Research and the US National Science Foundation, respectively.

Appendix A: Proof of Lemma 1

The proof is a direct consequence of the fact that the PPGLM likelihood described in our model has a global maximum with respect to its inputs [27]. To prove this, it is enough to show that the likelihoods in (7), (11) are concave functions of U. First for the log-link model, if we substitute (4) into (7) we have

$$L(\mathbf{N}|\mathbf{X}) = \sum_{c=1}^{C}\sum_{i=1}^{I}\left(\delta N_{c,i}\left(\theta_c^T\mathbf{x}_i + \log\Delta\right) - \Delta\exp\left(\theta_c^T\mathbf{x}_i\right)\right). \quad (41)$$

Stacking the difference process $\delta\mathbf{N}$ and the control input $\mathbf{U}$ into column vectors $\mathbf{n} \in \mathbb{R}^{CI}$, $\mathbf{u} \in \mathbb{R}^{SI}$ respectively and with modified parameter matrix $\overline{\Theta} \in \mathbb{R}^{CI\times(P+1)SI}$ corresponding to the extrinsic control part of the covariate matrix, we can write (41) as

$$L(\mathbf{N}|\mathbf{u}) = \mathbf{n}^T\overline{\Theta}\mathbf{D}'\mathbf{u} - \Delta\mathbf{1}^T\mathbf{K}\exp\left(\overline{\Theta}\mathbf{D}'\mathbf{u}\right) + \psi(\delta\mathbf{N})\log\Delta + r, \quad (42)$$

where $\mathbf{K} \in \mathbb{R}^{CI\times CI}$ is a diagonal matrix where the contributions of the process history and background activity for each process and time index are placed along the diagonals, $r = \mathbf{1}^T\log(\mathbf{K})\mathbf{n}$ is a constant (logarithm is applied to each element on the diagonal of $\mathbf{K}$) and $\mathbf{D}' \in \mathbb{R}^{(P+1)SI\times SI}$ is a design matrix that extracts the delayed inputs from $\mathbf{U}$ into a vector $\in \mathbb{R}^{(P+1)SI}$. Note that the likelihood is a combination of a linear term along with a negative exponential, which clearly makes the Hessian negative definite i.e. L is ***strictly*** concave with respect to the extrinsic inputs.

For the SEMPP with the logistic link model, we can similarly write (11)

$$L(\mathbf{N}^*|\mathbf{u}) = \mathbf{n}^{*T}\overline{\Theta}\mathbf{D}'\mathbf{u} - \mathbf{1}^T\log\left(1 + \mathbf{K}^s\exp\left(\overline{\Theta}\mathbf{D}'\mathbf{u}\right)\right) + r', \quad (43)$$

where $\mathbf{n}^* \in \mathbb{R}^{MI}$, $\mathbf{K}^s \in \mathbb{R}^{I\times MI}$ is a block diagonal matrix where each block is a row of contributions of the process history and background activity for each marked process and time index and $r' = \mathbf{1}^T\log(\mathrm{diag}(\mathbf{K}^s))\mathbf{n}$, $\mathrm{diag}(\mathbf{K}^s)$ is a diagonal matrix where the blocks of $\mathbf{K}^s$ constitutes the diagonal. Since we want to show the concavity of the likelihood with respect to $\mathbf{u}$, we can ignore the linear term in (43) and concentrate on the second term. Let us denote this as l_2,

$$l_2(\mathbf{u}) = -\sum_{i=1}^{I}\log\left(1 + \mathbf{k}_i^T\exp\left(\overline{\Theta}\mathbf{D}'\mathbf{u}\right)\right) \quad (44)$$

and $\mathbf{k}_i$ is the i-th row of the matrix $\mathbf{K}^s$. Note that all elements of $\mathbf{K}^s$ i.e. $k_{i,j} \geq 0$, $\forall i = 1,...,I$, $j = 1,...,MI$. Taking the gradient we have

$$\nabla l_2 = -\sum_{i=1}^{I}\frac{1}{1 + \mathbf{k}_i^T\exp(\overline{\mathbf{U}})}\Xi^T\mathrm{diag}\left(\mathbf{k}_i\right)\exp(\overline{\mathbf{U}}), \quad (45)$$

where $\Xi = \overline{\Theta}\mathbf{D}'$ and $\overline{\mathbf{U}} = \Xi\mathbf{u}$. Now let us denote $\mathrm{diag}\left(\mathbf{k}_i\right) = \mathbf{K}_i^s$ and calculate the Hessian for each i,

$$\nabla^2 l_2^i = -\frac{1}{z_i^2}(z_i\Xi^T\mathbf{K}_i^s\,\mathrm{diag}(\exp(\overline{\mathbf{U}}))\Xi \quad (46)$$
$$-\Xi^T\mathbf{K}_i^s\exp(\overline{\mathbf{U}})\exp(\overline{\mathbf{U}})^T\mathbf{K}_i^s\Xi),$$

where $z_i = 1 + \mathbf{k}_i^T\exp(\overline{\mathbf{U}})$. For strict concavity we need to show that $\forall\ \mathbf{y} \in \mathbb{R}^{MI}$,

$$\mathbf{y}^T\nabla^2 l_2\mathbf{y} < 0. \quad (47)$$

From (46) and (47),

$$\mathbf{y}^T\nabla^2 l_2^i\mathbf{y} = -\frac{1}{z_i^2}(\Xi\mathbf{y})^T\left(z_i\mathbf{K}_i^s\mathrm{diag}(\exp(\overline{\mathbf{U}})) - \mathbf{K}_i^s\exp(\overline{\mathbf{U}})\exp(\overline{\mathbf{U}})^T\mathbf{K}_i^s\right)\Xi\mathbf{y}. \quad (48)$$

Denoting $\Xi\mathrm{y} = \mathrm{w}$, $\overline{\mathbf{U}}^e = \mathrm{diag}(\exp(\overline{\mathbf{U}})$ and substituting z_i

$$\mathbf{y}^T \nabla^2 l_2^i \mathbf{y} = -\frac{1}{z_i^2}\left(\mathbf{w}^T \mathbf{K}_i^s \overline{\mathbf{U}}^e \mathbf{w} + \mathbf{w}^T \left(\sum_{j=1}^{MI} k_{i,j} \bar{u}_{j,j}^e\right) \mathbf{K}_i^s \overline{\mathbf{U}}^e \right. \quad (49)$$
$$\left. - \mathbf{K}_i^s \exp(\overline{\mathbf{U}}) \exp(\overline{\mathbf{U}})^T \mathbf{K}_i^s) \mathbf{w}\right)$$
$$= t_1^i + t_2^i .$$

Now let us analyze the second term separately,

$$t_2^i = -\frac{1}{z_i^2}\left(\sum_j k_{i,j} \bar{u}_{j,j}^e \sum_j k_{i,j} \bar{u}_{j,j}^e w_j^2 - \left(\sum_j k_{i,j} \bar{u}_{j,j}^e w_j\right)^2\right) \quad (50)$$
$$= -\frac{1}{z_i^2}\left(\sum_j v_{i,j} \sum_j v_{i,j} w_j^2 - \left(\sum_j v_{i,j} w_j\right)^2\right),$$
$$\text{where } v_{i,j} = k_{i,j} \bar{u}_{j,j}^e$$
$$\leq 0 \quad \left(\text{From Cauchy-Schwarz inequality}\right).$$

Now for the complete Hessian with (49), we have y

$$\mathbf{y}^T \nabla^2 l_2 \mathbf{y} = \sum_i \mathbf{y}^T \nabla^2 l_2^i \mathbf{y} = \sum_i \left(t_1^i + t_2^i\right) \quad (51)$$
$$\leq \sum_i t_1^i = -\mathbf{w}^T \sum_i \frac{1}{z_i^2} \mathbf{K}_i^s \overline{\mathbf{U}}^e \mathbf{w}$$
$$= -\mathbf{w}^T \overline{\mathbf{K}} \mathbf{w} < 0,$$

where $\overline{\mathbf{K}} = \sum_i \frac{1}{z_i^2} \mathbf{K}_i^s \overline{\mathbf{U}}^e \in \mathbb{R}^{MI \times MI}$ is a diagonal matrix and the negative definiteness comes from the fact that all the entries in matrix $\overline{\mathbf{K}}$ are positive, since the terms come from exponential of the co-variates. So the likelihood in (43) is ***strictly*** concave as well. Note that for any pattern **N** with at least one spike i.e., $\psi(\delta \mathbf{N}) > 0$, we can show

$$\nabla L(\mathbf{N}^* | \mathbf{u}) \not\to 0, \text{ if } \exists j \in \left\{1 \ldots SI\right\} \text{ such that } u_j \to \pm \infty . \quad (52)$$

Along with ***strict*** concavity, (52) means that the first-order condition for maximum is satisfied for a finite **U*** i.e., the maximum $L^* = L(\mathbf{N}|\mathbf{U}^*)$ will be global and unique. Now for

any ϵ with $\epsilon > L^*$, there is no control that satisfies (15) and thus ϵ- controllability is not achieved for unconstrained input.

Appendix B: Proof of Lemma 2

Consider an arbitrary realization $\mathbf{N}$ containing $\psi(\delta\mathbf{N})$ events overall and a box constraint on each extrinsic input i.e. $u_{s,i} \in \mathcal{U} = \left[u_{min}\ u_{max}\right]\ \forall\ s,i$. The likelihood in this case follows from (42)

$$L(\mathbf{N}|\mathbf{u}) = \mathbf{n}^T\overline{\Theta}\mathbf{D}'\mathbf{u} - \Delta\mathbf{1}^T\mathbf{K}\exp\left(\overline{\Theta}\mathbf{D}'\mathbf{u}\right) + \psi(\delta\mathbf{N})\log\Delta + r. \quad (53)$$

The first-order condition for maximum of $L(\boldsymbol{N} \mid \mathbf{u})$ is a transcendental equation, thus, the solution $\mathbf{U}^*$ cannot be derived in general and does not necessarily reside in the constrained space $\mathcal{U}$. Ideally we will need $L(\mathbf{U}^*)$ to analyze any dependence of maximum likelihood on spike count. But for $\mathbf{U} \in \mathcal{U}$ and bounded parameters Θ, the likelihood in (53) is dominated by the term $\psi(\delta\mathbf{N})\log\Delta$, i.e.,

$$\lim_{\Delta \to 0} L(\mathrm{N}|\mathbf{u}) \propto \frac{1}{\psi(\delta\mathbf{N})}\ \text{(since } \log\Delta < 0), \quad (54)$$

and a higher event count dictates the degradation of likelihood.

Appendix C: Proof of Lemma 3

Proof 5:

Here the variable of interest is the portion of the parameter vector Θ represented by $\beta_q^{m,c}$ for $q = 1...Q$, $m = 1...M$, $c = 1...C$ in (9). For the log link model in (7), $M = C$. Let us denote these set of values by $\alpha \in \mathbb{R}^{C^2Q}$ and rewrite the likelihood in (7) as a function of α following (42) in Appendix A,

$$L(\mathbf{N}|\alpha) = \mathbf{n}^T\mathbf{Z}\alpha - \Delta\mathbf{1}^T\mathbf{K}_p\exp(\mathbf{Z}\alpha) + \psi(\delta\mathbf{N})\log\Delta + r_\alpha, \quad (55)$$

where r_α is the contribution from other co-variates namely inputs and background activity independent of α, $\mathbf{Z} \in \mathbb{B}^{CI \times C^2Q}$ is a matrix composed of the relevant process history terms for each variable, time index and $\mathbf{K}_p \in \mathbb{R}^{CI \times CI}$ is a diagonal matrix similar to (42).

For the SEMPP model we can rewrite (11) for $\alpha \in \mathbb{R}^{MCQ}$ following (43),

$$L(\mathbf{N}^*|\alpha) = \mathbf{n}^{*T}\mathbf{Z}\alpha - \mathbf{1}^T\log\left(1 + \mathbf{K}_p^s\exp(\mathbf{Z}\alpha)\right) + r'_\alpha, \quad (56)$$

with $\mathbf{K}_p^s$ as the analog to $\mathbf{K}_p$ in (55). We note that both (55), (56) follow the same structure as their counterparts (42), (43) and thus we can conclude that the likelihoods are strictly concave with respect to the connectivity parameters a as well. Now to show that a critical amount of connectivity, e.g., a_c helps in the controllability of any arbitrary pattern, we investigate the first-order condition at $a = \mathbf{0}$. Computing the gradient of the likelihood in (55) we have

$$\begin{aligned}\nabla L_\alpha|_{\alpha=0} &= \left(\mathbf{n}^T\mathbf{Z}\right)^T - \Delta\mathbf{Z}^T\mathbf{K}_p\exp(\mathbf{Z}\alpha)|_{\alpha=0} \quad (57)\\ &= \mathbf{Z}^T\left(\mathbf{n} - \Delta\mathbf{K}_p\mathbf{1}\right).\end{aligned}$$

Now, the first-order condition for maximum is satisfied if

$$\mathbf{n} - \Delta\mathbf{K}_p\mathbf{1} \in \ker\left(\mathbf{Z}^T\right), \quad (58)$$

which does not hold in general for any $\mathbf{N}$, $\mathbf{U}$ and the rest of the parameters $\beta_0^c, \gamma_p^{m,s}\ \forall c,m,s$ and this proves that $a_c \neq 0$. We also claim that a_c does not diverge, i.e.,

$$\nabla L_\alpha|_{\alpha=\alpha_c} = \left(\mathbf{n}^T\mathbf{Z}\right)^T - \Delta\mathbf{Z}^T\mathbf{K}_p\exp\left(\mathbf{Z}\alpha_c\right) = 0 \quad (59)$$

has a solution. To see this, consider the case $\beta_1^{1,2} \to \infty$. Now since $\mathbf{Z} \in \mathbb{B}, \beta_1^{1,2} \to \infty$ implies a spike in the second neuron for previous time bin maximizes the probability of spike in the first neuron for the current time bin. But for a spike pattern $\mathbf{N}'$ in which such a sequence does not occur, the log-likelihood becomes

$$L\left(\mathbf{N}'|\alpha_c\right) \to -\infty. \quad (60)$$

If $\beta_1^{1,2} \to \infty$, likewise any pattern with consecutive spike from second and first neuron will have zero probability same as (60). This can also be seen from the first-order condition. So we can conclude that in general for any arbitrary pattern

$$\alpha_c \neq 0,\ |\alpha_c^j| \leq \alpha_{max} < \infty,\ \ \forall j = 1\ldots C^2Q. \quad (61)$$

Biography

Anirban Nandi (S'14) earned his B.E. from Jadavpur University in Kolkata, India, and is completing his Ph.D. in Electrical Engineering at Washington University in St. Louis. His primary research interests include control theory, dynamical systems, and computational neuroscience.

MohammadMehdi Kafashan (S'14-M'17) received the B.Sc. from Amirkabir University of Technology, the M.Sc. from Sharif University of Technology, and the Ph.D. from Washington University in St. Louis, all in Electrical Engineering. He currently holds a post-doctoral position with the Department of Neurobiology at Harvard Medical School. His research interests currently include computational neuroscience, statistical signal processing, and information processing in large neuronal networks.

ShiNung Ching (S'99-M'03) is currently an Assistant Professor in the Preston M. Green Department of Electrical and Systems Engineering at Washington University in St. Louis. He received his B. Eng (Hons.), M.A.Sc., and Ph.D., degrees from McGill University, the University of Toronto, and the University of Michigan, respectively. He was subsequently a postdoctoral associate at the Harvard Medical School. Dr. Ching's research interests are in the intersection of systems and control theory, theoretical neuroscience and neural engineering. He is author of over 60 refereed papers and the textbook Quasilinear Control.

References

[1]. Deisseroth K, "Optogenetics," Nature methods, vol. 8, no. 1, pp. 26–29, 2011. [PubMed: 21191368]

[2]. DiLorenzo PM and Victor JD, Spike timing: mechanisms and function. CRC Press, 2013.

[3]. Ching S and Ritt JT, "Control strategies for underactuated neural ensembles driven by optogenetic stimulation," Frontiers in neural circuits, vol. 7, 2013.

[4]. Nabi A, Mirzadeh M, Gibou F, and Moehlis J, "Minimum energy desynchronizing control for coupled neurons," Journal of computational neuroscience, vol. 34, no. 2, pp. 259–271, 2013. [PubMed: 22903565]

[5]. Iolov A, Ditlevsen S, and Longtin A, “Stochastic optimal control of single neuron spike trains,” Journal of neural engineering, vol. 11, no. 4, p. 046004, 2014. [PubMed: 24891497]

[6]. Nandi A, Ritt JT, and Ching S, “Non-negative inputs for underactuated control of spiking in coupled integrate-and-fire neurons,” in Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on IEEE, 2014, pp. 3041–3046.

[7]. Dasanayake I and Li J-S, “Optimal design of minimum-power stimuli for phase models of neuron oscillators,” Physical Review E, vol. 83, no. 6, p. 061916, 2011.

[8]. Boergers C and Kopell N, “Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity,” Neural Computation, vol. 15, no. 3, pp. 509–538, 2003. [PubMed: 12620157]

[9]. Snyder DL and Miller MI, Random point processes in time and space. Springer Science & Business Media, 2012.

[10]. McCullagh P and Nelder JA, Generalized linear models. CRC press, 1989, vol. 37.

[11]. Ogata Y, “Space-time point-process models for earthquake occurrences,” Annals of the Institute of Statistical Mathematics, vol. 50, no. 2, pp. 379–402, 1998.

[12]. Rodriguez-Iturbe I, Cox D, and Isham V, “Some models for rainfall based on stochastic point processes,” in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 410, no. 1839. The Royal Society, 1987, pp. 269–288.

[13]. Frost VS and Melamed B, “Traffic modeling for telecommunications networks,” Communications Magazine, IEEE, vol. 32, no. 3, pp. 70–81, 1994.

[14]. Paninski L, Pillow J, and Lewi J, “Statistical models for neural encoding, decoding, and optimal stimulus design,” Progress in brain research, vol. 165, pp. 493–507, 2007. [PubMed: 17925266]

[15]. Ahmadian Y, Packer AM, Yuste R, and Paninski L, “Designing optimal stimuli to control neuronal spike timing,” Journal of neurophysiology, vol. 106, no. 2, pp. 1038–1053, 2011. [PubMed: 21511704]

[16]. Dayan P and Abbott L, “Theoretical neuroscience: computational and mathematical modeling of neural systems,” Journal of Cognitive Neuroscience, vol. 15, no. 1, pp. 154–155, 2003.

[17]. Truccolo W, Eden UT, Fellows MR, Donoghue JP, and Brown EN, “A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects,” Journal of neurophysiology, vol. 93, no. 2, pp. 1074–1089, 2005. [PubMed: 15356183]

[18]. Okatan M, Wilson MA, and Brown EN, “Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity,” Neural computation, vol. 17, no. 9, pp. 1927–1961, 2005. [PubMed: 15992486]

[19]. Solo V, “Likelihood functions for multivariate point processes with coincidences.” in Decision and Control, 2007 46th IEEE Conference on IEEE, 2007, pp. 4245–4250.

[20]. Ba D, Temereanca S, and Brown EN, “Algorithms for the analysis of ensemble neural spiking activity using simultaneous-event multivariate point-process models,” Frontiers in computational neuroscience, vol. 8, 2014.

[21]. Daley DJ and Vere-Jones D, An introduction to the theory of point processes. Springer, 1988, vol. 2.

[22]. Whalen AJ, Brennan SN, Sauer TD, and Schiff SJ, “Observability and controllability of nonlinear networks: The role of symmetry,” Phys. Rev. X, vol. 5, p. 011005, 1 2015. [PubMed: 30443436]

[23]. Gerstner W and Kistler WM, Spiking neuron models: Single neurons, populations, plasticity. Cambridge university press, 2002.

[24]. Brown EN, Barbieri R, Ventura V, Kass RE, and Frank LM, “The time-rescaling theorem and its application to neural spike train data analysis,” Neural computation, vol. 14, no. 2, pp. 325–346, 2002. [PubMed: 11802915]

[25]. Kanamaru T, Horita T, and Okabe Y, “Theoretical analysis of arrayenhanced stochastic resonance in the diffusively coupled fitzhughnagumo equation,” Physical Review E, vol. 64, no. 3, p. 031908, 2001.

[26]. Izhikevich EM and FitzHugh R, “Fitzhugh-nagumo model,” Scholarpedia, vol. 1, no. 9, p. 1349, 2006.

[27]. Paninski L, “Maximum likelihood estimation of cascade point-process neural encoding models,” Network: Computation in Neural Systems, vol. 15, no. 4, pp. 243–262, 2004.

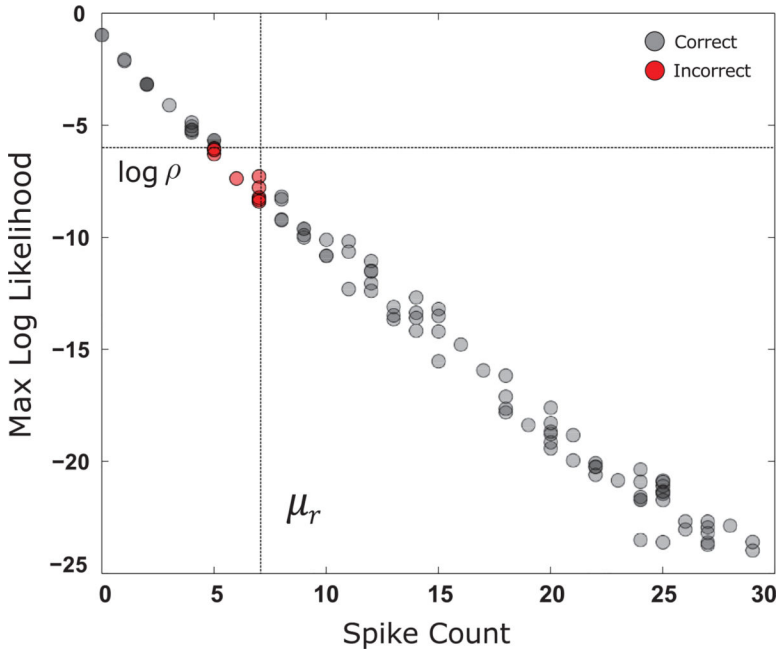

Fig. 1.

Accuracy of the relaxed maximally viable spike count (μ_r). The maximum likelihood is computed for 100 random realizations and compared to the predicted thresholds from the (μ_r,ρ)-controllability calculation ($\rho = 10^{-6}$).

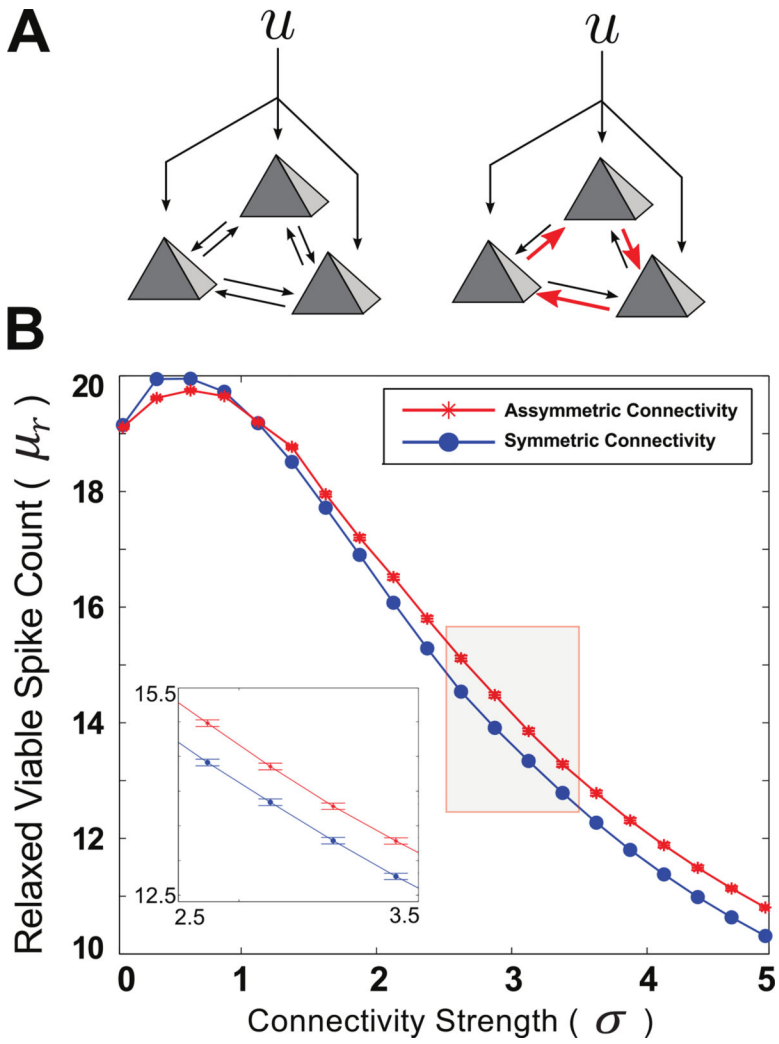

Fig. 2.

(A) Symmetric and Asymmetric 3-neuron motifs. (B) μ_r vs. Connectivity weight σ for $\rho = 10^{-10}$.

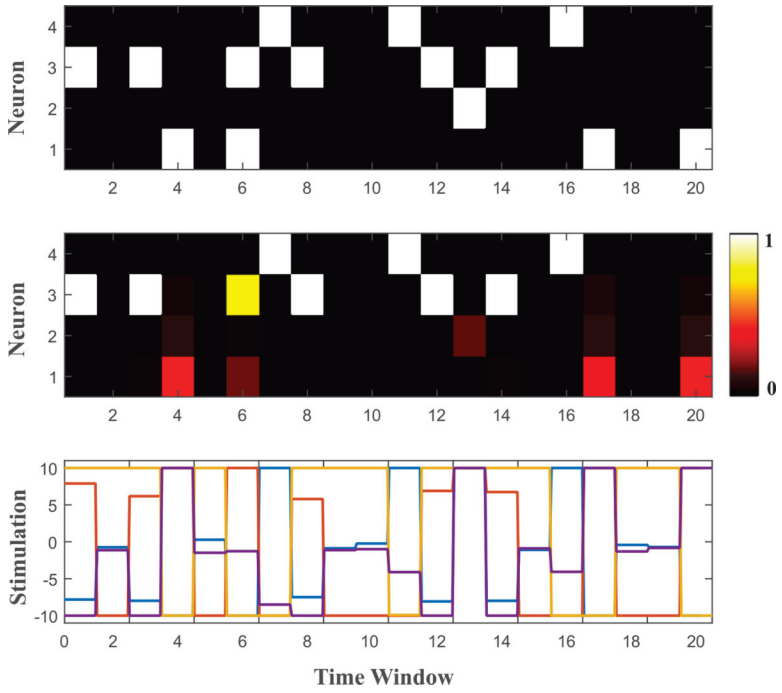

Fig. 3.

Validation of (μ, ρ)- viability. Top panel : A random pattern $\mathbf{N}T$ with $\psi(\delta\mathbf{N}T) < \mu r$ where μr is the solution of (33) for randomly chosen Θ. Middle panel : Achieved pattern in terms of probability of spiking in each window, with $\mathbf{U}^*$ from (36) and Bottom Panel : The optimized control U*from the MLE problem.

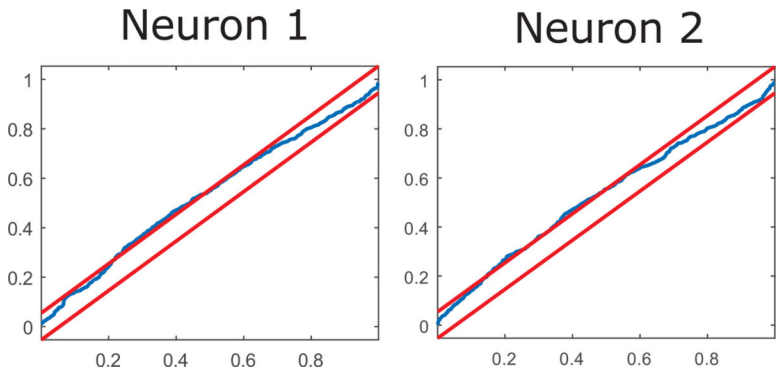

Fig. 4.

KS plots with 95% confidence bounds for goodness of fit assessment for a fully actuated two neuron INF network ($C = S = 2$) fitted using $Q = 3$ process history and $P = 5$ input history terms.

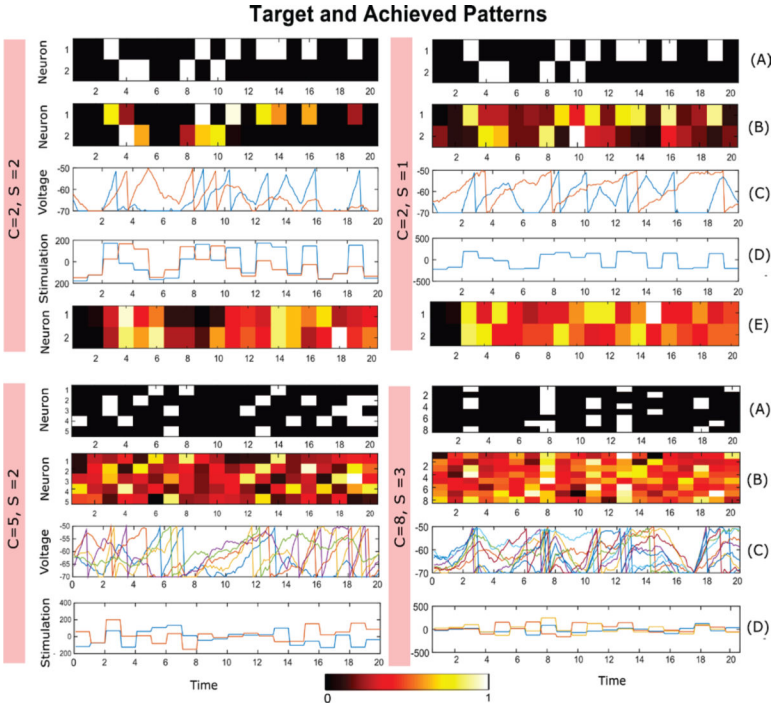

Fig. 5.

Control design for a *C* neuron coupled stochastic integrate and fire network. A : The target pattern for simulation study. B : The mean pattern over different realization for the INF neurons with the control **U***. C: One realization of voltage traces for the two INF neurons under **U***. D: The optimized input U*. E : The mean spiking pattern generated for a randomly selected **U** to validate optimality of **U***.

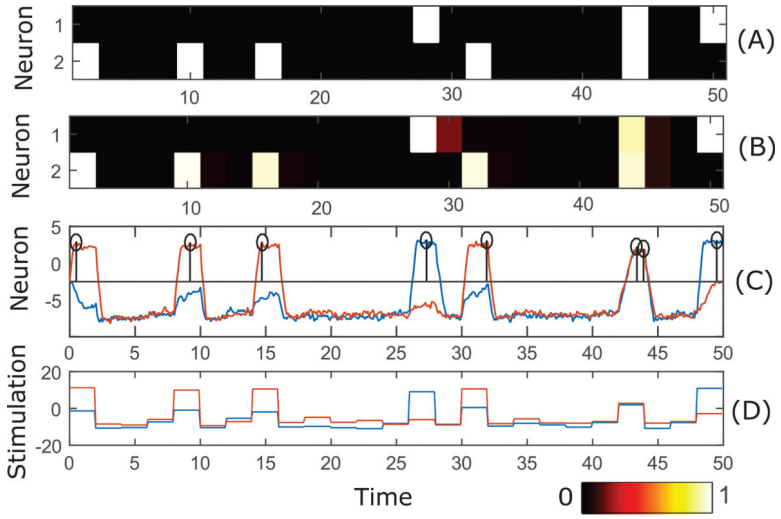

Fig. 6.

Control design for diffusively coupled FN neurons ($C = 2$). A: The target pattern $\mathbf{N}_T$ for the simulation. B: The average pattern over different realizations for the 2 neurons under **U***. C: One realization of voltage traces for the 2 FN neurons under **U***. The circled lines denote the detected spikes from the spike detection program. D: The optimized input U*.