# Behavior of integrate-and-fire and Hodgkin-Huxley models with correlated inputs

Jianfeng Feng
*COGS, Sussex University, Brighton, BN1 9QH, United Kingdom*

Ping Zhang
*St. Catharine College, Cambridge University, Cambridge CB2 1RL, United Kingdom*

(Received 13 September 2000; revised manuscript received 2 November 2000; published 9 April 2001)

We assess, both numerically and theoretically, how positively correlated Poisson inputs affect the output of the integrate-and-fire and Hodgkin-Huxley models. For the integrate-and-fire model the variability of efferent spike trains is an *increasing* function of input correlation, and of the ratio between inhibitory and excitatory inputs. Interestingly for the Hodgkin-Huxley model the variability of efferent spike trains is a *decreasing* function of input correlation, and for fixed input correlation it is almost independent of the ratio between inhibitory and excitatory inputs. In terms of the signal to noise ratio of efferent spike trains the integrate-and-fire model works better in an environment of asynchronous inputs, but the Hodgkin-Huxley model has an advantage for more synchronous (correlated) inputs. In conclusion the integrate-and-fire and Hodgkin-Huxley models respond to correlated inputs in totally opposite ways.

DOI: 10.1103/PhysRevE.63.051902 PACS number(s): 87.18.Sn

## I. INTRODUCTION

The two most commonly used single neuron models in theoretical neuroscience are the integrate-and-fire (IF) model, modeling neurons at an abstract level and the Hodgkin-Huxley (HH) model describing the biophysical mechanisms of cells. Recently it is found [1–3] that in certain parameter regions the coefficient of variation ($C_V$ = standard deviation/mean) of efferent spike trains of the HH model is almost independent of the ratio between inhibitory and excitatory inputs. In other words, whether inhibitory inputs are blocked or not has no effect on the $C_V$. However most results up date on the IF and HH models are obtained under the assumption that inputs are independent [1,4], both spatially and temporally. This assumption obviously contradicts the physiological data which clearly show that nearby neurons usually fire in a correlated way [5,6], and the anatomical data which reveal that neurons with similar functions group together and fire together. In fact ''firing together, coming together'' is a basic principle in neuronal development [7]. Furthermore, data in Ref. [5] indicate that even a weak correlation within a population of neurons can have a dramatic impact on the network behavior. The essential role played by the redundancy or correlation in perception has been appreciated early in the literature [8]. Hence it is of crucial importance to explore the impact of weakly correlated inputs on the efferent spike trains of neuronal models, which certainly sheds new light on the coding problem [9].

For the IF model the output firing variability is an increasing function of input correlation: the larger the input correlation is, the larger the $C_V$ of efferent spike trains. At the same time, as reported in many papers, the $C_V$ of efferent spike trains is an increasing function of the ratio between inhibitory and excitatory inputs. It is natural to expect that the behavior of the highly nonlinear HH model will be somewhat different from that of the IF model, which is a ''linear'' model *per se*. It is, however, surprising to find that the correlation in input signals has a totally *opposite* effect on the $C_V$ of efferent spike trains of the HH model: the correlation in inputs reduces rather than increases the $C_V$ of efferent spike trains. Furthermore, with fixed correlation, the $C_V$ of efferent spike trains is almost independent of the ratio between inhibitory and excitatory inputs. Hence the IF and HH models operate in two quite different modes: increasing the input correlation will decrease the signal to noise ratio of efferent spike trains in the IF model, whereas for the HH model an enhancement on the signal to noise ratio is attained. We thus conclude that the IF model works better in an environment of asynchronous inputs, but the HH model has an advantage for more synchronous (correlated) inputs. All the conclusions for the HH model are then repeated for the FitzHugh-Nagumo (FHN) model.

We then propose a simple approach called response surface to *graphically* explore the different behavior between two models. The response surface method enables us to grasp the property of a neuron with stochastic inputs. We advocate that the approach could also be applied in experiments.

Finally we employ the IF-FHN model [10] to show that the differences between the IF model and the HH model result from the fact that the former model has a constant decay rate, but the latter has a nonconstant decay rate. The IF-FHN model is obtained by extracting the leakage coefficient from the FHN model as exactly as possible. The leakage coefficient is a nonlinear function of $v$, in contrast to the constant leakage coefficient in the conventional integrate-and-fire model. The nonlinear leakage coefficient, taking a U-shape plotted against $v$, reveals interesting properties of the FHN model. The further the membrane potential is below its threshold, the stronger the leakage is and so the easier the model loses any memory of its recent history. Hence it is difficult to depolarize the cell when the membrane potential is far below the threshold. However when the membrane potential is near the threshold, the leakage gets smaller eventually becoming negligible; then incoming depolarizing signals can more easily induce the neuron to fire. In other

 

words, the neuron maintains its memory of recent activation in this range of membrane potential.

In conclusion, two most widely used neuron models, the integrate-and-fire and the Hodgkin-Huxley models with correlated inputs are considered. The irregularity of output spike trains of the integrate-and-fire model is an increasing function of input correlations, but is a decreasing function for the Hodgkin-Huxley model. A graphic approach, the response surface method, is proposed to intuitively reveal the differences between the ouput activity of two models. Using the IF-FHN model, we theoretically explore the underpining mechanisms of the models and conclude that an appropriate form of nonlinear decay rate can account for the aforementioned differences between the integrate-and-fire and the Hodgkin-Huxley models.

The paper is organized as follows. In Sec. II the integrate-and-fire model and the Hodgkin-Huxley model are introduced. In Sec. III correlated synaptic inputs are defined. Section IV is devoted to numerical results of the integrate-and-fire model, the Hodgkin-Huxley model, and the FitzHugh-Nagumo model with correlated inputs. In Sec. V the response surface approach is introduced and applied to the integrate-and-fire model and the Hodgkin-Huxley model. Finally in Sec. VI, the IF-FHN model is briefly reviewed [10] and its behavior with correlated inputs is shown. In the present paper, we exclusively restrict ourselves to the models with Poisson, positively correlated inputs and refer the reader to Ref. [11] for the negatively correlated input case.

## II. THE INTEGRATE-AND-FIRE MODEL AND HODGKIN-HUXLEY MODEL

Suppose that a cell receives excitatory postsynaptic potentials (EPSPs) at $p$ synapses and inhibitory postsynaptic potentials (IPSPs) at $q$ inhibitory synapses. The activities among excitatory synapses and inhibitory synapses are correlated but, for simplicity of notation only, are assumed to be independent between them. When the membrane potential $V_t$ is between the resting potential $V_{\rm rest}$ and the threshold $V_{\rm thre}$, it is given by

$$dV_t = -\frac{1}{\gamma}(V_t - V_{\rm rest})dt + d\overline{I}_{\rm syn}(t), \tag{2.1}$$

where $1/\gamma$ is the decay rate and synaptic inputs

$$\overline{I}_{\rm syn}(t) = a\sum_{i=1}^{p} E_i(t) - b\sum_{j=1}^{q} I_j(t)$$

with $E_i(t), I_i(t)$ as Poisson processes with rate $\lambda_E$ and $\lambda_I$, respectively, $a>0, b>0$ being the magnitude of each EPSP and IPSP. Once $V_t$ crosses $V_{\rm thre}$ from below a spike is generated and $V_t$ is reset to $V_{\rm rest}$. This model is termed as the IF model. The interspike interval (ISI) of efferent spikes is

$$T = \inf\{t: V_t \geqslant V_{\rm thre}\}$$

and the $C_V$ of interspike intervals (ISIs) is given by

$$C_V = \sqrt{\langle T^2\rangle - \langle T\rangle^2}/\langle T\rangle.$$

For the simplicity of notation we assume that the correlation coefficient between the $i$th excitatory (inhibitory) synapse and the $j$th excitatory (inhibitory) synapse is $c(i,j) = \rho(||i-j||)$, where $\rho$ is a nonincreasing function. A slight more general model than the IF model defined above is the IF model with reversal potentials defined by

$$dZ_t = -\frac{Z_t - V_{\rm rest}}{\gamma}dt + d\overline{I}_{\rm syn}(Z_t, t), \tag{2.2}$$

where

$$\overline{I}_{\rm syn}(Z_t, t) = \overline{a}(V_E - Z_t)\sum_{i=1}^{p} E_i(t) + \overline{b}(V_I - Z_t)\sum_{j=1}^{q} I_j(t).$$

$V_E$ and $V_I$ are the reversal potentials $V_I < V_{\rm rest} < V_E$, $\overline{a}(V_E - V_{\rm rest}), \overline{b}(V_I - V_{\rm rest})$ are the magnitude of single EPSP and IPSP when $Z_t = V_{\rm rest}$. $Z_t$ (membrane potential) is now a birth-and-death process with boundaries $V_E$ and $V_I$.

We consider the classic HH model with correlated inputs given by

$$\begin{aligned} CdV = &-g_{Na}m^3h(V - V_{Na})dt - g_k n^4(V - V_k)dt \\ &- g_L(V - V_L)dt + dI_{\rm syn}(V,t), \end{aligned} \tag{2.3}$$

where $I_{\rm syn}(V,t) = \overline{I}_{\rm syn}(t)$ or $I_{\rm syn}(V,t) = \overline{I}_{\rm syn}(V,t)$. Equations and parameters used in the HH model are as follows (see Refs. [1,12]):

$$\frac{dn}{dt} = \frac{n_\infty - n}{\tau_n}, \quad \frac{dm}{dt} = \frac{m_\infty - m}{\tau_m}, \quad \frac{dh}{dt} = \frac{h_\infty - h}{\tau_h},$$

and

$$n_\infty = \frac{\alpha_n}{\alpha_n + \beta_n}, \quad m_\infty = \frac{\alpha_m}{\alpha_m + \beta_m}, \quad h_\infty = \frac{\alpha_h}{\alpha_h + \beta_h},$$

$$\tau_n = \frac{1}{\alpha_n + \beta_n}, \quad \tau_m = \frac{1}{\alpha_m + \beta_m}, \quad \tau_h = \frac{1}{\alpha_h + \beta_h},$$

with

$$\alpha_n = \frac{0.01(V+55)}{\exp\left(-\frac{V+55}{10}\right) - 1}, \quad \beta_n = 0.125\exp\left(\frac{V+65}{80}\right),$$

$$\alpha_m = \frac{0.1(V+40)}{\exp\left(-\frac{V+40}{10}\right) - 1}, \quad \beta_m = 4\exp\left(\frac{V+65}{18}\right),$$

$$\alpha_h = 0.07\exp\left(-\frac{V+65}{20}\right), \quad \beta_h = \frac{1}{\exp\left(-\frac{V+35}{10}\right) + 1}.$$

The parameters used in Eq. (2.3) are $C=1, g_{Na}=120, g_K=36, g_L=0.3, V_k=-77, V_{Na}=50$, and $V_L=-54.4$.

### III. SYNAPTIC INPUTS

Here we use the usual approximation to approximate the IF models with or without reversal potentials, or more exactly the synaptic inputs of the models. We do not check the approximation accuracy since it has been done by many authors [4,13].

The input now reads

$$E_i(t) \sim \lambda_E t + \sqrt{\lambda_E} B_i^E(t)$$

and similarly

$$I_i(t) \sim \lambda_I t + \sqrt{\lambda_I} B_i^I(t),$$

where $B_i^E(t)$ and $B_i^I(t)$ are standard Brownian motions. Therefore the IF model without reversal potentials can be approximated by

$$dv_t = -\frac{1}{\gamma}(v_t - V_{\rm rest})dt + d\bar{i}_{\rm syn}(t),$$

where

$$\bar{i}_{\rm syn}(t) = a\sum_{i=1}^{p} \lambda_E t - b\sum_{j=1}^{q} \lambda_I t + a\sqrt{\lambda_E}\sum_{i=1}^{p} B_i^E(t) - b\sqrt{\lambda_I}\sum_{j=1}^{q} B_j^I(t). \tag{3.1}$$

Since the summation of Brownian motions is again a Brownian motion we can rewrite the equation above as

$$\bar{i}_{\rm syn}(t) = \mu t + \sigma B(t), \tag{3.2}$$

where $B(t)$ is a standard Brownian motion

$$\begin{cases} \mu = ap\lambda_E - bq\lambda_I, \\ \sigma^2 = a^2 p\lambda_E + b^2 q\lambda_I + a^2\lambda_E \sum_{i\neq j}^{p} c(i,j) + b^2\lambda_I \sum_{i\neq j}^{q} c(i,j). \end{cases} \tag{3.3}$$

The input variance $\sigma^2$ is an increasing function of $c_{i,j}$.

Now we turn our attention to the IF model with reversal potentials. Similar to what we have done for the IF model without reversal potentials, we can rewrite down the model in the following form:

$$dz_t = -\frac{1}{\gamma} z_t dt + d\bar{i}_{\rm syn}(z_t, t), \tag{3.4}$$

where

$$\bar{i}_{\rm syn}(z_t, t) = \{ap[V_E - z(t)]\lambda_E - bq[z(t) - V_I]\lambda_I\}t + \sigma(z_t)B(t)$$

and

$$\sigma^2(z_t) = a^2(z_t - V_e)^2 p\lambda_E + b^2(z_t - V_I)^2 q\lambda_I + a^2(z_t - V_E)^2\lambda_E \sum_{i\neq j}^{p} c(i,j) + b^2(z_t - V_I)^2\lambda_I \sum_{i\neq j}^{q} c(i,j). \tag{3.5}$$

There are other forms of the diffusion terms in Eq. (3.4) to approximate the original process [14]. For simplicity of notation we confine ourselves to $\sigma^2(z_t)$.

For the HH model we have analogous expressions for correlated inputs with or without reversal potentials as in Eq. (3.2), i.e., $\bar{i}_{\rm syn}(t)$ and Eq. (3.4), i.e., $\bar{i}_{\rm syn}(v,t)$. In the sequel, we confine ourselves to the case of inputs without reversal potentials. Nevertheless all results are qualitatively true for the models with inputs with reversal potentials (not shown, see Ref. [2]).

### IV. NUMERICAL RESULTS

From the results above we conclude that correlated inputs increase the variance of inputs and thus we expect an increase of $C_V$ of efferent interspike intervals for the IF model [3,15]. For the HH model an analytical treatment is difficult and we have to resort to numerical simulations. Here only results of the case that $c(i,j) = c$ for $i \neq j$, for the models without reversal potentials are presented. It has been reported in the literature that the correlation coefficient between cells is around 0.1 in V5 of a rhesus monkeys *in vivo* [5], and around 0.2 in human motor units of a variety of muscles [16].

Figure 1 shows numerical results of the IF model and HH model with $c$ between 0 and 0.1, and $q = 0,10,\cdots,100$. We first look at the results for the IF model [Fig. 1(a)]. When $c = 0$, inputs without correlation, there are many numerical and theoretical investigations (see, for example, Refs. [1–3,10,11,17]). In line with the theoretical results above, the larger the input correlation is, the larger the output $C_V$. In particular we note that when $c > 0.08$ we have $C_V > 0.5$ for any value of $q$. In conclusion, with a fixed $p$, $C_V$ is an increasing function of the correlation and the number of inhibitory inputs. However for the HH model the situation is totally different [Fig. 1(b)]: for fixed correlation coefficients $C_V$ is almost a constant, independent of the number of inhibitory inputs; for fixed number of inhibitory inputs $C_V$ is a *decreasing* function of the input correlation. In other words,

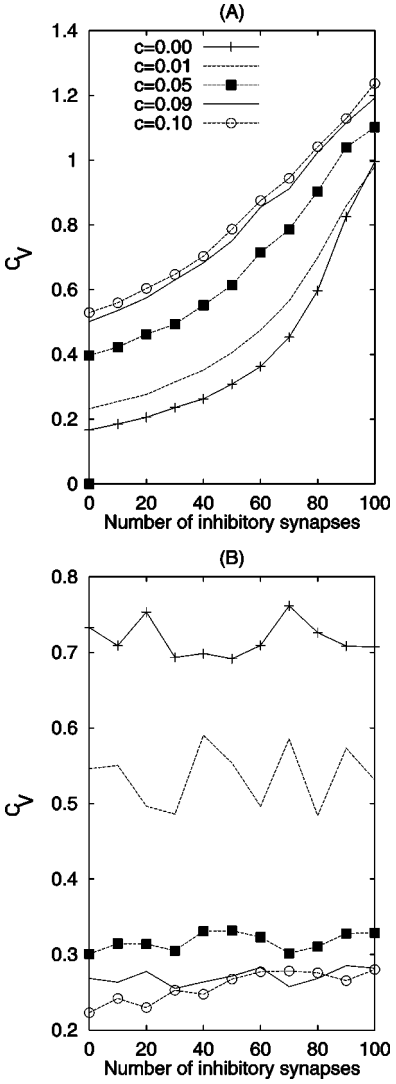


FIG. 1. $C_V$ vs correlation and $q$. Data are obtained by simulating the IF model (a) and the HH model (b) with synaptic input $\overline{i}(t)$. $a=b=0.5$ mV, $\lambda_I=\lambda_E=100$ Hz, $V_{\rm rest}=0$ mV, $V_{\rm thre}=20$ mV, $\gamma=20.2$ msec, $p=100$, $q=0,10,\ldots,100$. For the reason of the choice of these parameters we refer the reader to Ref. [1].

for HH model the stronger the input correlation is (equivalent to a stronger input noise), the more regular the output.

In recent years there has been much research devoted to the problem that how to generate efferent spikes with a $C_V$ between 0.5 and 1 (see, for example, Refs. [1–3,10,11,15,17,18]). In particular it is pointed out in Ref. [17], where only independent inputs $c=0$ are considered, that it is impossible for the IF model and some biophysical models to generate spikes with a $C_V$ between 0.5 and 1 if the inputs are exclusively excitatory. This phenomenon is referred to as ''central limit effect'' and widely cited in the literature (see, for example, Ref. [19]). Many different approaches have been proposed to get around this problem. In the present paper we clearly demonstrate that even with exclusively excitatory inputs the IF model is capable of emitting spike trains with a $C_V$ greater than 0.5. Under the condition that $c>0.08$ [Fig. 1(a)], no matter what the ratio between the excitatory and inhibitory synapses is, the $C_V$ of efferent spike trains is always greater than 0.5. For the HH model when the input correlation is low, independent of the inhibitory inputs, the output $C_V$ is always greater than 0.5.

There is increasing evidence to support the assumption that the brain might use different coding strategies when dealing with tasks of different complexity. When a slow reaction is required, the brain might use rate coding to process information. When a fast reaction is needed the brain only has enough time to take into account the first spike of active synapses and so temporal coding seems more plausible. This assumption naturally require that neurons operate in different modes. Here we provide such an example of different operation modes, provided that both the IF and the HH mechanisms are employed by cells in the brain. The IF model is sensitive to its input correlation and so it would work in an environment of less correlated or asynchronous inputs. The HH model is more reliable when correlated or synchronized inputs are presented. Figure 2 (see Fig. 4) shows a case of signal to noise ratio of efferent spike trains, i.e., mean/standard deviation or $1/C_V$. When input correlation is small the $S_{NR}$ of the IF model is greater than that of the HH model, but when it is larger, the $S_{NR}$ of the IF model is less than that of the HH model.

Note that the above properties of $S_{NR}$ are true for all parameter regions we considered (see next section): no matter if its $C_V$ is greater than 0.5 or not. Hence when we restrict ourselves to the parameter regions of irregular firing: for the HH model with a fixed number of inhibitory inputs, the $S_{NR}$ is an *increasing* function of the input correlation; for the IF model with a fixed number of inhibitory inputs, the $S_{NR}$ is a *decreasing* function of the input correlation.

An obvious difference between the integrate-and-fire model and the HH model lies in the fact that the latter one has an refractory period of about 12.2 msec [1]. We also add refractory periods to the integrate-and-fire model for calculating related quantities and all results above remain true (see next section). In fact, from Fig. 2 we conclude that adding a refractory period to the IF model will even increase the discrepancy between two models. According to the definition of $S_{NR}$ we know that $\overline{S_{NR}}=(\langle T\rangle+R)/S_D=S_{NR}+R/S_D$, where $\overline{S_{NR}}$ is obtained after adding a refractory period $R$, and $S_D$ is the standard deviation of interspike intervals of the IF model. Since $S_D$ is an increasing function of correlations, we see that $\overline{S_{NR}}$ will more sharply decrease than the $S_{NR}$ of the IF model shown in Fig. 2, and results in even large differences between two models.

Essentially it is impossible to have an analytical treatment of the HH model. For confirmation of our results on the HH model, we also simulated the HH model in NEURON [20] with synaptic inputs as a square wave of magnitude of 5 $\mu$ Amp, duration 0.1 msec [1]. The results match approximately with Fig. 1(b). For both models, we have carried out systematic simulations on a wide range of parameters ranging from $p=75,100,150,200$ and $0\leqslant r=q/p\leqslant 1$. The results obtained qualitatively agree with the conclusions above.

We also simulate another simplified model, the FitzHugh-Nagumo model which mimics the HH model, with correlated

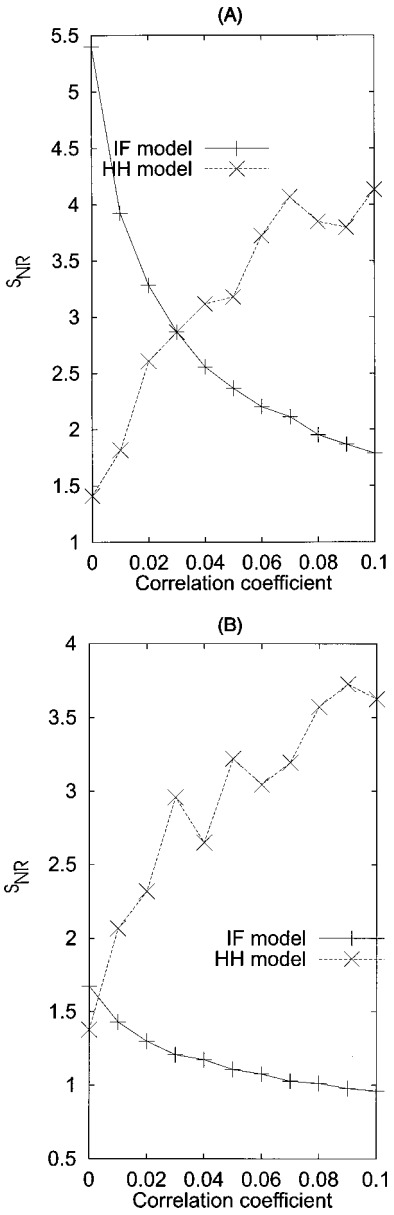


FIG. 2. Signal to noise ratio ($S_{NR}$) of efferent spike trains of the IF model and HH model, comparing Fig. 1. Left figure corresponds to the case of $1/C_V$ with fixing $q=10$ in Fig. 1; right figure with fixing $q=80$.

inputs. Figure 3 shows that all conclusions above for the HH model are true for the FitzHugh-Nagumo model. The FitzHugh-Nagumo model is described by

$$\begin{cases} dv=\gamma[-v(v-\alpha)(v-1)-w]dt+d\overline{i}_{\rm syn}, \\ dw=\delta[v-\beta w]dt \end{cases} \tag{4.1}$$

with $\alpha=0.2,\beta=2.5,\gamma=100,\delta=0.25,\ a=b=0.06$ in $\overline{i}_{\rm syn}$ and all other parameters are the same as in the HH model.

All numerical simulations for the IF model are carried out with synaptic inputs defined by Eq. (3.2) with a time step 0.01 and a Euler scheme [21]. For the HH model, an algorithm for solving stiff equations from NAG library is used with step size 0.01. Further small time steps are used and we conclude no significant improvements are obtained. When calculating mean firing rate and $C_V$, 10 000 interspike intervals are employed.

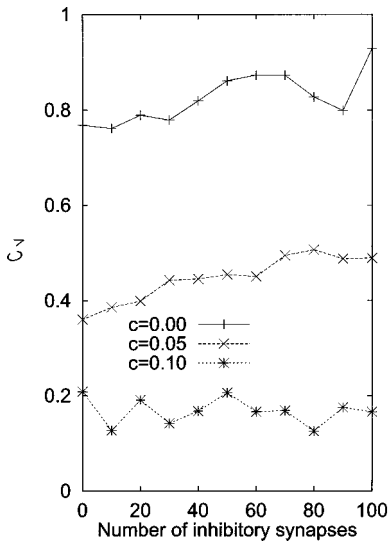


FIG. 3. $C_V$ vs q for the FitzHugh-Nagumo model with correlated inputs $p=100,\lambda_E=\lambda_I=100$ Hz.

## V. RESPONSE SURFACES

Neurons have traditionally been characterized by the nature of their so-called *F-I* curve, that is the relationship between the rate of firing *F*, that they adopt in response to an applied current and the level of the applied current *I*, proposed by Hodgkin. He classified membranes as type *I*, if they can show an arbitrarily low firing rate and long spike latency in response to a continuous current; or type II, if they exhibit a narrow range of response firing rates (not close to zero), and virtually zero spike latency. The HH model is classified as type II. In a sense, the IF model can be classified as type I, since arbitrarily low firing rates are possible for just suprathreshold currents. This is a useful categorization for many purposes. Yet frequently neurons are subject to input regimes which cannot be approximated as a constant current, as we consider here. Inputs often take a pulse form, and, in numerous brain areas, e.g., the visual cortex and the hypothalamus, neurons fire apparently randomly. For a neuron model with stochastic inputs, the *F-I* curve only provides us with limited information.

We thus propose a simple characterization of neuronal response to random synaptic input. This involves a graphical presentation of the first two moments, the mean and variance, of neuronal output as function of the first two moments of total synaptic inputs, for a range of values of these input moments. The usual approximations of a wide variety of stochastic input process are first constructed using the usual diffusion approximation [see Eq. (3.3)]. The measures of mean and variability of output we use are conventional measures: overall firing rate $F$ and $C_V$. Hence when a neuron model receives inputs ranging from exactly balanced inputs

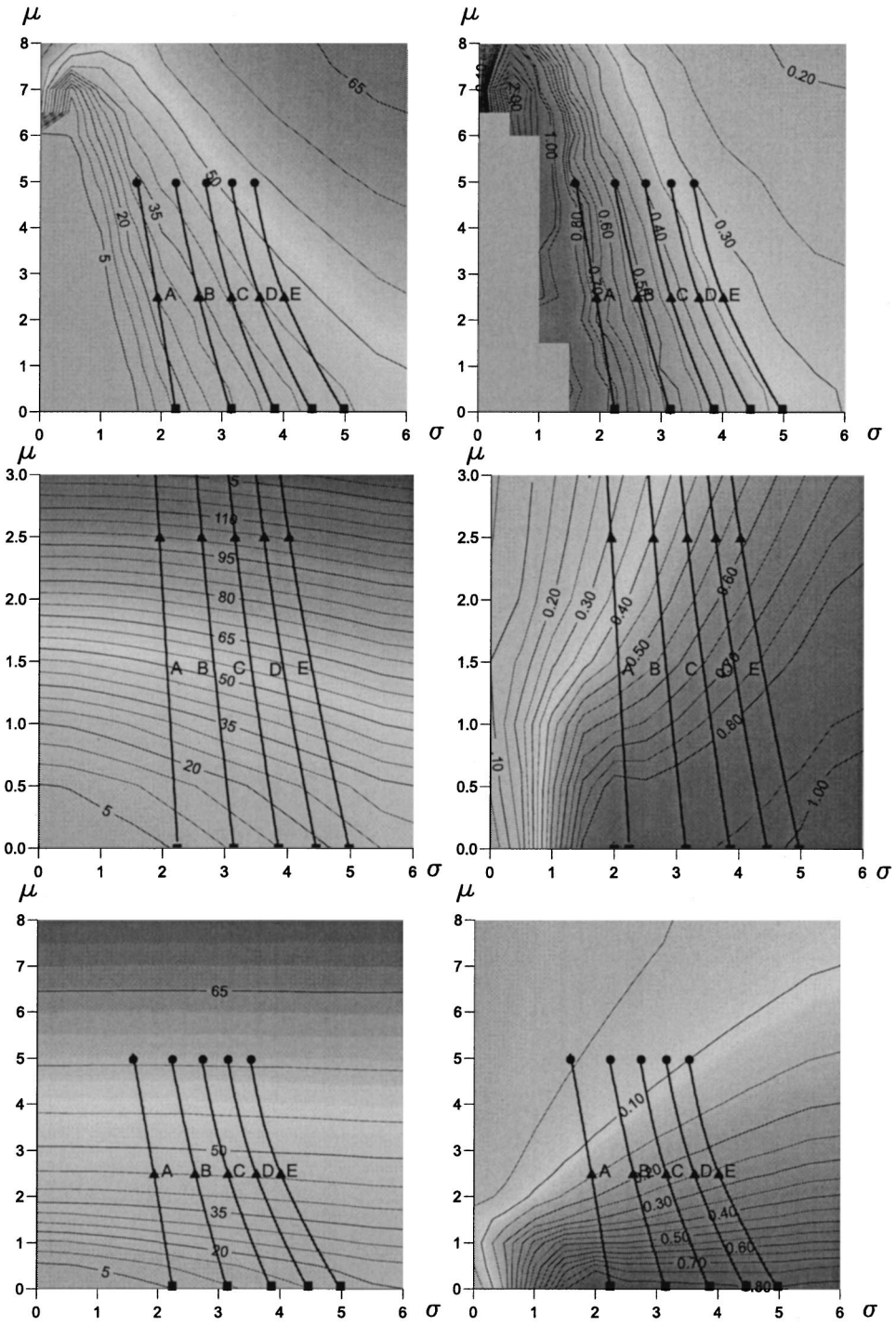


FIG. 4. Response surfaces of the HH and IF model with correlated inputs $c=0.00$ (a), $c=0.01$ (b), $c=0.02$ (c), $c=0.03$ (d), $c=0.04$ (e). Top panel: left is $F-(\mu,\sigma)$, right is $C_V$-$(\mu,\sigma)$ for the HH model; middle panel: left is $F-(\mu,\sigma)$, right is $C_V$-$(\mu,\sigma)$ for the IF model without a refractory period; bottom panel: left is $F-(\mu,\sigma)$, right is $C_V$-$(\mu,\sigma)$ for the IF model after adding a refractory period of 12 msec, which is approximately the refractory period of the HH model [1].

to purely excitatory inputs, its behavior could be fully understood by simply looking up the trajectory on the surface.

Applying the approach to the HH model and the IF model yields response surfaces as depicted in Fig. 4. Lines[1] (a)–(e) are trajectories for efferent firing frequency and $C_V$, when $\mu$ goes from 0 (exactly balanced input) to $a\lambda_E N_E$ (purely excitatory input) or equivalently $r$ from 0 to 1. We could easily tell the different response behavior of the two models. Lines (a)–(e) cross contours of the $C_V$-$(\mu,\sigma)$ surface of the IF model, but they are almost in parallel with contours of the HH model. Comparison of (a)–(e) in Fig. 4 demonstrates the huge changes in the noise component, $\sigma$, as a result of introducing very low correlation. The increase is greatest changing from $c=0$ to $c=0.01$, and, for the HH model, the consequent increase in $F$ is biggest. At balance $(\mu=0)$, $F$ increases from 16 to 29 Hz, whereas for pure excitation input, the change is from 34 to 43 Hz, a rather smaller increase. Each subsequent equal sized increment in correlation induces lower increases in $F$ (e.g., to 37,41,44 Hz at balance)

[1]All parameters are the same as in the previous section, see Fig. 1.

for two reasons: the increment in $\sigma$ is much lower at higher correlations[2] and the relationship between $F$ and $\sigma$ has a lower slope at higher values of $\sigma$. However, for different values of $p$ and $a$, and therefore different starting points on the $\mu,\sigma$ surface for independent inputs, the qualitative pattern of correlation-induced increments in $F$ might be very different. For the IF model, positive correlation increases $F$ but by a much smaller amount (by virtually zero for purely excitatory input for the IF model). The effect on $C_V$ is the opposite to that for the HH model: it increases variability rather than reduces it, over much of the range. We have applied the approach to analyze other interesting behaviors of the models and will report it in a further publication.

## VI. IF-FHN MODEL

In order to theoretically understand the results presented in the previous sections, we consider the IF-FHN model, which mimics the FHN model and is proposed in Ref. [10], with correlated inputs. First we briefly review the model.

The basic idea to derive the IF-FHN model is that to develop a systematic approach to approximating biophysical models by models of the integrate-and-fire type. The two essential components of the leaky integrate-and-fire model are integration of incoming signals and leakage. Our approach is then to determine terms which reflect these two components for a given biophysical model as exactly as possible. Devising methods for approximating biophysical models by abstract models — which preserve the essential complexity of the biophysical mechanism yet are simultaneously concise and transparent — is an important continuing task in computational neuroscience. The advantages are obvious. Biophysical models are usually difficult to understand, and to simulate at a network level, characteristics not shared, for example, by the conventional integrate-and-fire (IF) model. A simplified expression might also provide us with a new tool to understand the frequently puzzling behavior of biophysical models, since the response of the conventional leaky integrate-and-fire type model to stochastic input is more comprehensible. Although a rigorous analytical treatment is difficult, various approximations are available (see, for example, Ref. [13]).

As an application of the idea above, we consider the FitzHugh-Nagumo (FHN) model. We first define leakage coefficient as precisely as possible in this more general context. Consider a general model

$$\begin{cases} dv(t)=f(v,w)dt, \\ dw(t)=g(v,w)dt \end{cases} \tag{6.1}$$

in which $v$ is membrane potential, $w$ is a vector recovery variable, generally representing activation and inactivation variables for the ion channels in the model. Our aim in achieving an integrate-and-fire reduction is to re express the model as

$$dv(t)=-L(v)(v-v_{\text{rest}})dt, \tag{6.2}$$

where $v_{\text{rest}}$ is the resting potential and $L(v)$ is the generalized leakage coefficient. In the extreme case, for the conventional leaky integrate and fire model, the model is extremely simple

$$dv(t)=-L(v-v_{\text{rest}})dt, \tag{6.3}$$

where the leakage coefficient $L=1/\gamma$ is a constant, independent of the values of $v$. In general, it will only be possible to express a model in the form of Eq. (6.2) approximately.

Now we mimic the FHN model with leaky integrate-and-fire type model. First of all we see that the second differential equation of the FHN model can be solved as

$$w(t)=\delta\int_0^t v(s)\exp[-\beta\delta(t-s)]ds. \tag{6.4}$$

Substituting Eq. (6.4) into the first differential equation of the FHN model we obtain

$$dv(t)=-\gamma(v-1)(v-\alpha)vdt-\delta\int_0^t v(s) \times\exp[-\beta\delta(t-s)]ds+d\bar{i}_{\text{syn}}(t). \tag{6.5}$$

Using our basic idea, to extract the leakage coefficient from the FHN model as exactly as possible, i.e., to rewrite this differential equation in the form

$$dv(t)=-L(v)(v-v_{\text{rest}})dt, \tag{6.6}$$

we rewrite Eq. (6.5) as follows

$$\begin{aligned} dv(t)=&-\left[\gamma(v-1)(v-\alpha)+\delta\int_0^t \times\exp[-\beta\delta(t-s)]ds\right]v(t)dt-\delta\int_0^t[v(s) \\ &-v(t)]\exp[-\beta\delta(t-s)]dsdt+d\bar{i}_{\text{syn}}(t) \\ =&-\left[\gamma(v-1)(v-\alpha)+\frac{1}{\beta}[1-\exp(-\beta\delta t)]\right]v(t)dt \\ &-\delta\int_0^t[v(s)-v(t)]\exp[-\beta\delta(t-s)]dsdt \\ &+d\bar{i}_{\text{syn}}(t). \end{aligned} \tag{6.7}$$

Note that in the equation above the term

$$\delta\int_0^t[v(s)-v(t)]\exp[-\beta\delta(t-s)]ds$$

[2]Remembering that $\sigma^2(c)=a_1+ca_2$ where $a_1,a_2$ are positive constants defined in Eq. (3.3), we have $\sigma^2(c+c_1)-\sigma^2(c)=a_2c_1$ which is independent of $c$, but $\sigma(c+c_1)-\sigma(c)=\sqrt{a_1+(c+c_1)a_2}-\sqrt{a_1+ca_2}$ which is a decreasing function of $c$ for fixed $c_1>0$.

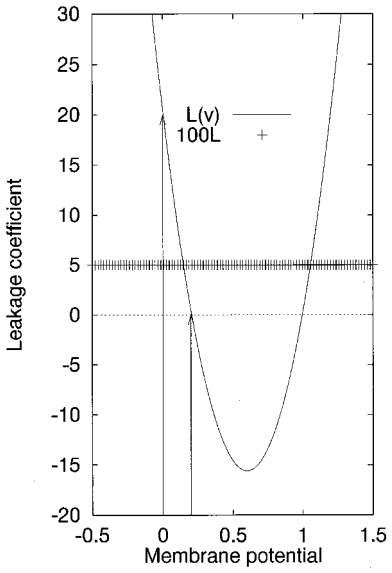


FIG. 5. $L(v)$ and $100L=5.0$ vs the membrane potentials. Between the threshold and the resting potential (indicated by arrows), $L(v)$ are positive, but the closer the membrane potential to the threshold, the weaker the leakage, in contrast to the conventional integrate-and-fire model.

is a higher order term and which we could omit in the first order approximation. Equation (6.7) becomes

$$dv(t)=-\left[\gamma(v-1)(v-\alpha)+\frac{1}{\beta}[1-\exp(-\beta\delta t)]\right]v(t)dt + d\bar{i}_{\mathrm{syn}}(t). \tag{6.8}$$

Let us define

$$L(v)=\gamma(v-1)(v-\alpha)+\frac{1}{\beta} \tag{6.9}$$

which gives us the leakage coefficient (approximated to the first order) extracting from the FHN model.

Figure 5 depicts a typical case of the leakage coefficient extracted from the FHN model. When the membrane potential is between the resting potential $V_{\mathrm{rest}}\sim 0$ and the threshold $V_{\mathrm{th}}\sim\alpha$(indicated by arrow), the leakage coefficient is positive. Hence the system will gradually lose its memory of recent activation. However, $L(v)$ is very different from $L$, which is a constant and is independent of its membrane potentials. $L(v)$ is larger when the membrane potential is close to the resting potential, and vanishes when the membrane potential is close to the threshold. In other words, when the membrane potential is near resting potential, the model loses its memory rapidly. Incoming signals accumulate less effectively to increase membrane potential. When membrane potential is near to the threshold, however, the FHN model behaves more similar to a perfect integrate-and-fire model. The FHN now has a very good ''memory'' and in a sense ''waits'' just below the threshold. As soon as some positive signals arrive, the neuron fires. Therefore, below the threshold, the IF-FHN behaves as a combination of the *leaky* integrate-and-fire model and the *perfect* integrate-and-fire model.

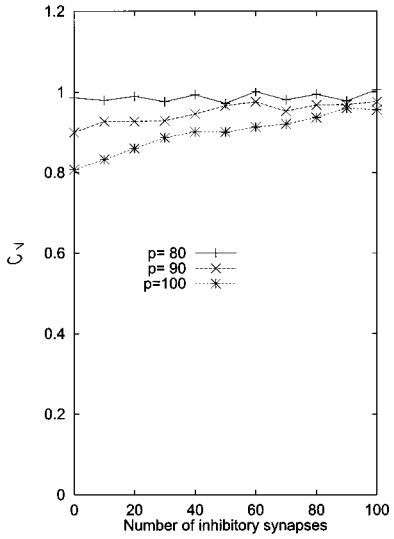


FIG. 6. $C_V$ vs $q$ and standard deviation of interspike intervals vs mean interspike intervals of the IF-FHN model. With respect to varying $q$, $C_V$ (calculated after adding a refractory period of 3.2 msec) is quite flat. Note that the standard deviation equals almost equals the mean interspike interval.

Once the membrane potential is above the threshold, now $L(v)$ acts as an amplifier of incoming signal, rather than as a leakage. It will increase membrane potential until it arrives at its maximum value, designated as $v_h$ in this paper, and then $L(v)$ becomes positive again.

Now we are in the position to define the following dynamics as the integrate-and-fire model with nonlinear leakage (IF-FHN) [10]:

$$\begin{cases} dv_t=-L(v)v\,dt+d\bar{i}_{\mathrm{syn}}(t), \\ v_0=v_{\mathrm{rest}}. \end{cases} \tag{6.10}$$

For a prefixed *value* $v_h$ (defined before), once $v$ crosses it from below, $v$ is then reset to $v_{\mathrm{rest}}$. Unlike the conventional integrate-and-fire model, the IF-FHN increases to $v_h$ rather than $V_{\mathrm{th}}$, which is smaller than $v_h$.

We use the set of parameters as before. What is the behavior of the IF-FHN? In Fig. 6 we see that $C_V$ is quite high and is not sensitive to the number of inhibitory inputs, similar to what we have observed for the FHN model itself [1]. However, in Ref. [1] we were not able to elucidate the mechanism which ensures the occurrence of the phenomenon. Based upon the numerical results on the IF-FHN model, we conclude that the nonlinear leakage coefficient contributes to the flat $C_V$ which is a typical feature of some biophysical models and which is not captured by conventional integrate-and-fire models.

To further demonstrate the power of our approach, we consider the models with correlated inputs. The IF model is essentially a linear model and so the larger is its input fluctuation, the larger is its output variety. But biophysical models such as the FHN and Hodgkin-Huxley models exhibit

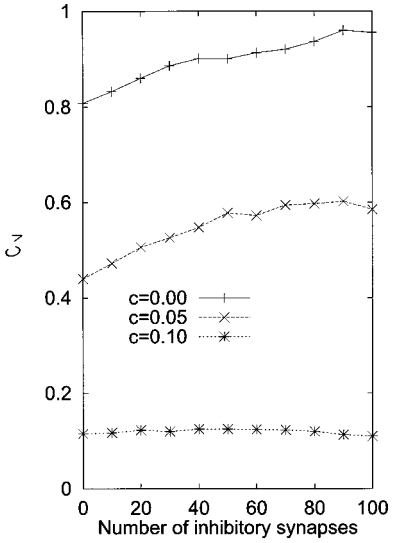


FIG. 7. $C_V$ vs $q$ of the IF-FHN model. With fixed $q$, $C_V$ is a decreasing function of correlation $c$; in other words, the signal-to-noise ratio ($S_{NR}$), $S_{NR}=1/C_V$, is an increasing function of $c$, $p=100$. Compare with Fig. 3.

opposite behavior. This phenomenon is of particular interest. Biophysical models such as the FHN and the Hodgkin-Huxley models improve their performance in an environment of correlated inputs which is certainly of extremely possible case in a neuronal assembly. The phenomenon is also the main motivation of our present approach in this section *per se*: to find integrate-and-fire type models which are qualitatively in agreement with biophysical models. Suppose that the correlation between synapses is $c>0$. Figure 7 clearly shows now the contradictory between biophysical models and integrate-and-fire type model is resolved. Our simulations further confirm that the key difference between the FHN model and the conventional integrated-and-fire model lies in the fact that the former has a nonlinear leakage coefficient. More specifically, let us look at the $C_V$-$(\mu,\sigma)$ surface of the IF-FHN model. First, $C_V$ is a decreasing function of the input correlation coefficient $c$ (Fig. 7), as in the HH model (Fig. 4). Secondly, for inputs with a fixed $C$, when $r$ varies from 0 to 1, $C_V$ changes, but only slightly. For example, when $c=0.1$, $C_V$ of the IF-FHN model is close to 0.1, almost independent of $r$. This fact implies that $C_V$ follows the contours of the response surface, as in the HH model (Fig. 4).

From the data shown in Fig. 6 we might envisage that Kramer's formula can predict the model behavior, which has been successfully applied to estimate the firing rate in certain circumstances [22]. Kramer's formula (a special case of the large deviation theory, see Ref. [23], and references therein for details [24]) reads

$$\langle \overline{T} \rangle \sim \frac{2\pi}{\sqrt{H''(v_{\min})|H''(v_{\max})|}} \exp[2(H_{\max}-H_{\min})/\sigma^2], \tag{6.11}$$

where $\overline{T}$ is the first exit time from a potential well, and $H_{\max}$,$H_{\min}$ are the local maximum and local minimum of the potential well (see below).

Furthermore $\overline{T}$ is exponentially distributed (similar to the interspike intervals in a Poisson process), as the plot of standard deviation vs interspike intervals of Fig. 6 shows, i.e. standard deviation equals mean. If this is the case then we might conclude that the flat $C_V$ is simply due to perturbations of a deterministic system, a simple, classic and clear picture. Denote

$$\overline{H}(v)=\int_0^v uL(u)du=\gamma\left(\frac{1}{4}v^4-(\alpha+1)\frac{1}{3}v^3+\frac{1}{2}\alpha v^2\right)+\frac{1}{2\beta}v^2.$$

For the IF-FHN model we could write the potential $H$ of the system in terms of two terms

$$H(v)=\overline{H}+\mu v \tag{6.12}$$

and therefore

$$H_{\max}-H_{\min}=\overline{H}(v_{\max})-\overline{H}(v_{\min})+\mu(v_{\max}-v_{\min}),$$

where $v_{\max}$ and $v_{\min}$ are the value at which $H$ attains the local maximum and minimum. As we mentioned before for the IF-FHN model, its behavior does not substantially change if we set the threshold as a value inside $[v_{\max},1]$.

Figure 8 shows an application of Kramer's formula to the IF-FHN model with correlated inputs. When the input is uncorrelated $c=0.0$, Kramer's formula gives a rough estimate, with an obvious discrepancy between numerical results and theoretical estimate. Nevertheless when a small correlation is added ($c\geqslant 0.005$), i.e., the IF-FHN model receives a more random input, Kramer's formula gives an excellent estimate. As one might expect, the mean interspike interval and standard deviation exhibits a linear relationship.

## VII. DISCUSSION

In summary, in Ref. [1] independent inputs to neuronal models are considered and it is reported that the IF model and the HH model behave *differently*. The present work carries this a considerable stage further in that, when correlated inputs are considered, the IF and HH model behave in totally *opposite* ways: the $S_{NR}$ of the IF model decreases with the increase of the input correlation; while the $S_{NR}$ of the HH model increases. The conclusions of the present work would be informative, in particular when we deal with network models where inputs to each unit are bound to be correlated. For a given neuron with stochastic inputs, the response surface provides us with a valuable way to understand its behavior. Finally, using the IF-FHN model, we pointed out that one of the key difference between the IF model and the HH

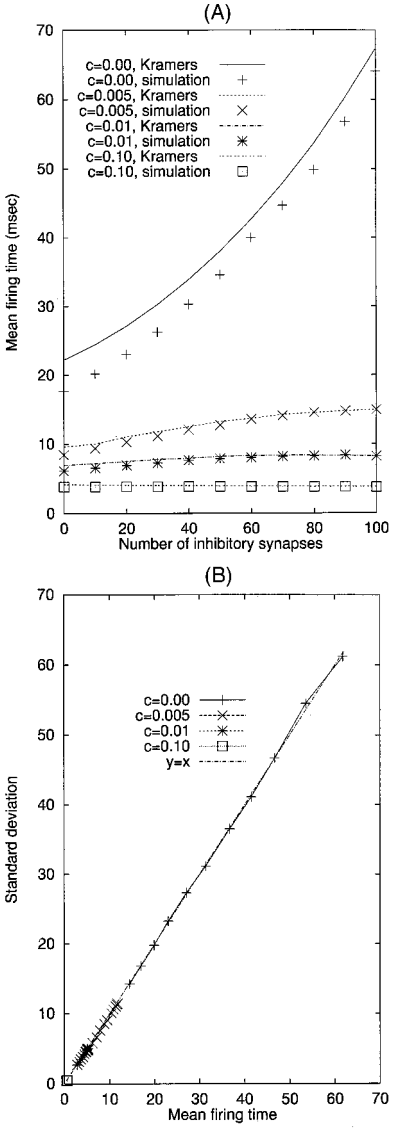


FIG. 8. A comparison between Kramer's formula and numerical simulations with parameters as specified in the context, $p=100$.

model is that the leakage coefficient is a constant for the former, but it depends on the membrane potential for the latter. The difference between the IF model and the IF-FHN model might also provide a clue for the phenomenon observed in Ref. [25].

Let us now discuss some related issues which we have not taken into account in the present paper.

We have only considered the case of so-called spatially correlated inputs. What would happen when temporal correlations are also taken into account? Here we have to first distinguish two kinds of temporal correlations: the correlation between spikes of a single input spike train and the correlation due to the rising and decay time of input spikes (noninstantaneous inputs such as the $\alpha$-synapse input [22]). For the former case, we could easily approximate inputs by the diffusion approximation, as in Sec. III. To understand the model behavior is then simply reduced to looking up the response surface of the model. For the latter case, it is more complicated. The diffusion approximation will not be valid if the rising or decay time is long enough, which means that, instead of the Brownian motion, we have to resort to the Ornstein-Uhlenbeck process to approximate the synaptic input. A thorough consideration of the latter case is outside the scope of the present paper and we will report it in further publications.

In the present paper, we assume that there are correlations among excitatory synapses and inhibitory synapses, but not between them. It is also illuminating to consider the case that (a) inhibitory inputs are totally independent or (b) inhibitory inputs and excitatory inputs are correlated with a correlation coefficient $c^{E,I}(i,j)$. For the case (a), $\sigma^2$ in Eq. (3.3) becomes

$$\sigma^2 = a^2 p \lambda_E + b^2 q \lambda_I + a^2 \lambda_E \sum_{i \neq j}^{p} c(i,j). \tag{7.1}$$

Therefore in the $C_V$-$(\mu,\sigma)$ surface (see Fig. 4), all trajectories (b)–(e) with $c>0$ and $q>0$ ($q=0$ corresponds to the points indicated by circle) have a shift towards the left. For example, when $c=0.04$ and $q=p$ the point indicated by the rectangle of the trajectory $E$ is at point $\sqrt{14.9}$ rather than 5. In words, the model behavior is less affected by input correlations. For the case (b), similarly we have

$$\sigma^2 = a^2 p \lambda_E + b^2 q \lambda_I + a^2 \lambda_E \sum_{i \neq j}^{p} c(i,j) + b^2 \lambda_I \sum_{i \neq j}^{q} c(i,j) - 2ab\sqrt{\lambda_E \lambda_I} \sum_{i=1}^{p} \sum_{j=1}^{q} c^{E,I}(i,j) \tag{7.2}$$

Depending on whether inhibitory and excitatory inputs are positively or negatively correlated, the input variance $\sigma^2$ could be either reduced or enhanced. Nevertheless, the model behavior could be understood by a simple calculation of $\mu,\sigma$ and looking up the response surfaces.

For a ''linear'' model such as the IF model, we naturally expect that an increase in the input mean or variance results in an increase in the output mean or variance. For some nonlinear models such as the HH model and the IF-FHN model, our results tell us that it is not always the case. In fact it is well known in the literature that the HH model can, in certain circumstances, increase its firing rate with an increase of inhibitory inputs, the so-called post inhibitory rebound [26]. Our results here reveal another interesting, similar behavior of the HH model: an increase in the input variability result in a decrease in the output variability.

## ACKNOWLEDGMENTS

We thank David Brown, Guibin Li, and Andrew Davison and for their comments on early versions of this manuscript. The work was partially supported by BBSRC and an ESEP grant of the Royal Society.

[1] D. Brown, J. Feng, and S. Feerick, Phys. Rev. Lett. **82**, 4731 (1999); J. Feng, *ibid.* **79**, 4505 (1997); J. Feng, and D. Brown, J. Phys. A **31**, 1239 (1998); J. Feng and D. Brown, Biol. Cybern. **78**, 369 (1998); J. Feng, D. Brown, and G. Li, Phys. Rev. E **61**, 2987 (2000).
[2] J. Feng and D. Brown, Biol. Cybern. **80**, 291 (1999).
[3] J. Feng and D. Brown, Neural Comput. **12**, 711 (2000).
[4] H.C. Tuckwell, *Stochastic Processes in the Neurosciences* (Society for Industrial and Applied Mathematics, Philadelphia, 1988).
[5] E. Zohary, M.N. Shadlen, and W.T. Newsome, Nature (London) **370**, 140 (1994).
[6] C.F. Stevens and A.M. Zador, Nature Neurosci. **1**, 210 (1998).
[7] B.R. Sheth J. Sharma, S.C. Rao, and M. Sur, Science **274**, 2110 (1996).
[8] H. Barlow, in *Functions of the Brain*, edited by C. Coen (Clarendon Press, Oxford, 1986), pp. 11–43.
[9] T.D. Albright, T.M. Jessell, E.R. Kandel, and M.I. Posner, Cell **100**, s1 (2000); M.N. Shadlen and J.A. Movshon, Neuron **24**, 67 (1999); J. Feng and R. Cassia-Moura, Phys. Rev. E **59**, 7246 (1999).
[10] J. Feng and D. Brown, Bull. Math. Biol. **62**, 467 (2000).
[11] J. Feng and B. Tirozzi, Phys. Rev. E **61**, 4207 (2000).
[12] L.F. Abbott, in *Neural Modeling and Neural Networks* edited by F. Ventriglia (Pergamon Press, Oxford, 1994), pp. 57–78.
[13] L.M. Ricciardi and S. Sato, *Lectures in Applied Mathematics and Informatics*, edited by L.M. Ricciardi (Manchester University Press, Manchester, 1990).
[14] M. Musila and P. Lánský, J. Theor. Biol. **171**, 225 (1994).
[15] M.N. Shadlen and W.T. Newsome, Curr. Opin. Neurobiol. **4**, 569 (1994); M.N. Shadlen and W.T. Newsome, J. Neurosci. **18**, 3870 (1998).
[16] P.B.C. Mattews, J. Physiol. (London) **492**, 597 (1996).
[17] W. Softky and C. Koch, J. Neurosci. **13**, 334 (1993).
[18] P. Konig, A.K. Engel, and W. Singer, Trends Neurosci. **19**, 130 (1996).
[19] L.F. Abbott, J.A. Varela, K. Sen, and S.B. Nelson, Science **275**, 220 (1997).
[20] M.L. Hines and N.T. Carnevale, Neural Comput. **9**, 1179 (1997).
[21] J. Feng, G. Lei, and M. Qian, J. Comput. Math. **10**, 376 (1992).
[22] C. Koch, *Biophysics of Computation* (Oxford University Press, Oxford, 1999).
[23] S. Albeverio, S. Feng, and M. Qian, Phys. Rev. E **52**, 6593 (1995).
[24] S. Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1989).
[25] R. Azouz and C.M. Gray, Proc. Natl. Acad. Sci. U.S.A. **197**, 8110 (2000).
[26] N. Kopell and G. LeMasson, Proc. Natl. Acad. Sci. U.S.A. **91**, 10 586 (1990).