Increasing inhibitory input *increases* neuronal firing rate: why and when? Diffusion process cases

Increasing inhibitory input *increases* neuronal firing rate: why and when? Diffusion process cases

Jianfeng Feng[1] and Gang Wei[2]

[1] COGS, Sussex University, BN1 9QH, UK
[2] Department of Mathematics, Hong Kong Baptist University, Hong Kong, People's Republic of China

E-mail: jf218@cam.ac.uk and gwei@math.hkbu.edu.hk

Abstract
Increasing inhibitory input to single neuronal models, such as the FitzHugh–Nagumo model and the Hodgkin–Huxley model, can sometimes increase their firing rates, a phenomenon which we term inhibition-boosted firing (IBF). Here we consider neuronal models with diffusion approximation inputs, i.e. they share the identical first- and second-order statistics of the corresponding Poisson process inputs. Using the integrate-and-fire model and the IF-FHN model, we explore theoretically how and when IBF can happen. For both models, it is shown that there is a critical input frequency at which the efferent firing rate is identical when the neuron receives purely excitatory inputs or exactly balanced inhibitory and excitatory inputs. When the input frequency is lower than the critical frequency, IBF occurs.

PACS number: 87.19.La

1. Introduction

The reason why neurons in the cortex receive and emit stochastic rather than deterministic signals remains elusive, despite a century of research activity. The advantage of deterministic signal transmission over stochastic is obvious: it is much more economic and reliable. The stochastic part of a signal is usually thought of as 'noise' and is the part any system tries to get rid of. In stochastic resonance theory [5, 19], noise is hypothesized to be useful, but an application of the theory to the neuronal system tells us that it only works inside a very limited parameter region and that a carefully adjusted input signal is required [14]. Another possible role played by noise in neuronal systems is that, with the help of noise, a neuron, such as the Hodgkin–Huxley [20, 31] and the FitzHugh–Nagumo [17, 24, 31] model (not shown), can *increase* its efferent firing rate when inhibitory inputs *increase*. As in the literature [4, 7, 12], we

assume that a neuron receives inputs ranging from purely excitatory inputs to exactly balanced inhibitory and excitatory inputs.

A natural and interesting question is then why and when increasing inhibitory inputs to a neuron can boost its efferent firing rate. A full treatment of the Hodgkin–Huxley and FitzHugh–Nagumo model with stochastic inputs is difficult, although this might not always remain 'a formidable task' [31]. The traditional and theoretically tractable way to deal with the models with Poisson inputs is to consider the models with diffusion approximation inputs, i.e. inputs with identical mean and variance as the Poisson process. We first numerically show that the Hodgkin–Huxley model exhibits the inhibition-boosted firing (IBF) when it receives diffusion approximation inputs. We then turn to the integrate-and-fire (IF) model and the IF-FHN model with diffusion approximation inputs. The IBF phenomenon is observable for both models, which indicates that IBF is not due to complex, nonlinear mechanisms of biophysical models. Theoretical results for the IF model and IF-FHN model are developed, which elucidates the mechanism underpinning the IBF phenomenon.

The IF model is one of the most widely used models in (theoretical) neuroscience. It is an extremely simplified model and is linear (before resetting). Nevertheless, in the literature, a transparent and theoretical result on its input–output relationship is lacking (see for example, equation (9.238) in [31]). We apply theoretical results to the IF model and a simple relationship between input and output frequency is obtained. The formula enables us to prove that when the input frequency is lower than a critical frequency, the output frequency is higher when the neuron receives a mixture of inhibitory and excitatory inputs than when it receives purely excitatory inputs. Moreover, the critical frequency is unique. For the IF-FHN model, a model which is originally proposed to mimic the FitzHugh–Nagumo model, we apply Kramer's formula to prove that there is a critical frequency at which the efferent firing rate when the model receives purely excitatory inputs is equal to the rate with exactly balanced inputs.

Roughly speaking, the IBF phenomenon is due to a competition between two driving forces of neurons: stochastic and deterministic. Assume that a is the magnitude of EPSPs (excitatory postsynaptic potentials) or IPSPs (inhibitory postsynaptic potentials) and λ is the input frequency. When the neuron receives purely excitatory inputs, the deterministic force is proportional to $a\lambda$ and the stochastic force is $a^2\lambda$. For the exactly balanced input case, the deterministic force is 0 and stochastic force is $2a^2\lambda$. In general, the deterministic force is more efficient in driving a cell to fire and therefore increasing inhibitory input reduces the firing rate of a neuron, in agreement with our intuition. However, when λ is small enough, the deterministic force of purely excitatory inputs is $a\lambda$ and the deterministic force plays a minor role in driving the cell to fire. In other words, now the noise term is more prominent. The noise term for the exactly balanced input case is $2a^2\lambda$, which is twice that for purely excitatory inputs, $a^2\lambda$. Therefore under these circumstances the neuron fires faster when inhibitory inputs increase, i.e. noise increases.

The above scenario provides us with the answer to the 'why' question. It is of equal importance to answer the 'when' question, since in parameter regions where the IBF phenomenon occurs the neuron might fire too slowly and have no physiological reality. For the IF model and in parameter regions used in the literature, this is truly the case. It is difficult to observe it if only numerical simulations are employed. This might also tell us that why the IBF phenomenon has never been reported in the literature. Nevertheless, for the IF-FHN model, there is a physiologically reasonable region of (a, λ) in which increasing inhibitory inputs increases neuronal firing rate, as we have observed for the Hodgkin–Huxley model and the FitzHugh–Nagumo model. We fully characterize this region for the IF-FHN model. As we have pointed out before [16], the nonlinear leakage in the IF-FHN model ensures that it behaves very differently from the IF model.

The arguments above also indicate that increasing inhibitory inputs boosting neuronal firing rate is a universal phenomenon. Whether we can observe it or not in a physiologically plausible parameter region depends on neuronal parameters, or for real neurons, on the environment in which they operate. Since a neuron usually receives a massive excitatory and inhibitory input, we hope our finding may shed new light onto the coding problem [15, 18] and suggest another functional role of inhibitory inputs [32], or noise terms in signal inputs.

As a by-product, our results also alter another conventional view in theoretical neuroscience: increasing inhibitory inputs results in an increase of the randomness of output spike trains. In recent years, the issue has been extensively discussed [1, 4, 7, 27, 28]. A more general and biologically realistic principle is that the faster a neuron fires, the more regular its interspike intervals are. In terms of this principle, we demonstrate that in the parameter regions in which the IBF occurs, the efferent interspike intervals become more regular when inhibitory inputs increase. Here the regularity of efferent spike trains is measured by the coefficient of variation (CV) of interspike intervals, i.e. standard deviation/(mean + refractory period) for the IF and the IF-FHN model.

2. Models

For two given quantities $V_{\text{thre}} > V_{\text{rest}}$ and when $v_t < V_{\text{thre}}$, the membrane potential v_t satisfies the following dynamics:

$$\mathrm{d}v_t = -L(v_t)v_t\,\mathrm{d}t + \mathrm{d}I_{\text{syn}}(t)$$
$$v_0 = V_{\text{rest}}. \tag{2.1}$$

$I_{\text{syn}}(t)$ is the synaptic input given by

$$\mathrm{d}I_{\text{syn}}(t) = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}B_t \tag{2.2}$$

with constants $\mu \geqslant 0, \sigma \geqslant 0$ and the standard Brownian motion B_t [31]. Once v_t is greater than V_{thre}, it is reset to V_{rest}. More specifically we define

$$\mu = a\lambda(1-r) \qquad \sigma^2 = a^2\lambda(1+r) \tag{2.3}$$

where $a > 0$ is the magnitude of EPSPs and IPSPs, λ is the input rate and r is the ratio between inhibitory inputs and excitatory inputs. In particular, when $r = 0$ the neuron exclusively receives excitatory inputs; when $r = 1$ the inhibitory and excitatory input is exactly balanced. Here, for simplicity of notation, we assume that the EPSP and IPSP size are equal and refer the reader to [1, 22] for a more complete and biologically oriented formulation of synaptic inputs.

When $L(v) = L > 0$, a constant, the model is termed the IF model; when the leakage coefficient

$$L(v) = \gamma(v - V_{\text{thre}})(v - \alpha) + \frac{1}{\beta} \tag{2.4}$$

with constants $\gamma > 0, \alpha > 0, \beta > 0$, the model is as proposed in [10] for mimicking the FitzHugh–Nagumo model.

Unlike the IF model, for the IF-FHN model the leakage coefficient is not a constant. When the membrane potential is between the resting potential $V_{\text{rest}} \sim 0$ and the threshold V_{thre}, the leakage coefficient $L(v)$ is positive. Hence the system will gradually lose its memory of recent activation. However, $L(v)$ is very different from L. $L(v)$ is larger when the membrane potential is close to the resting potential. The more the membrane potential is away from its threshold, the larger the decay is. This mechanism naturally prevents the membrane becoming too negative, a mechanism the IF model lacks, and we have to set a lower bound for the IF model. The leakage coefficient vanishes when the membrane potential is close to the threshold. In
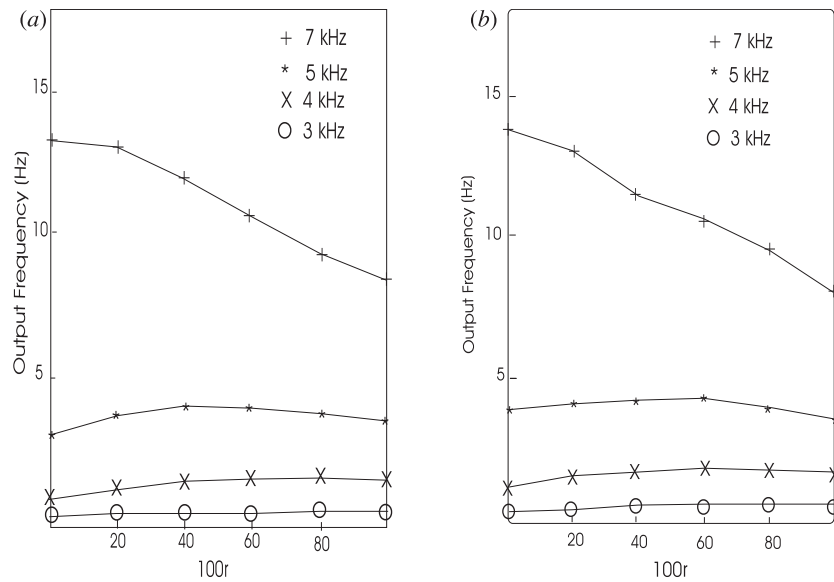
Figure 1. Mean output frequencies versus $100r$ for the HH model with $\lambda = 7, 5, 4, 3$ kHz, $a = 0.5$ mV (*a*) diffusion approximation inputs, (*b*) corresponding Poisson inputs.

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 IF-FHN model behaves more like a perfect IF model. The IF-FHN model 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 model behaves as a combination of the *leaky* IF model and the *perfect* IF model.

Once the membrane potential is above a certain value, $L(v) < 0$ and now it acts as an amplifier of incoming signal, rather than as a leakage. It will increase membrane potential until it arrives at its maximum value and then $L(v)$ becomes positive again.

In the sequence, we define

$$T(r) = \inf\{t : v_t \geqslant V_{\text{thre}}\} \tag{2.5}$$

as the mean firing time for $r \in [0, 1]$. For the IF-FHN model a constant refractory period of $T_{\text{ref}} = 3.2$ ms is usually added to $T(r)$, which is approximately the refractory period of the FitzHugh–Nagumo model [4].

3. Examples

We first consider the classic HH model with synaptic inputs given by

$$C \, dV = -g_{Na} m^3 h(V - V_{Na}) \, dt - g_k n^4 (V - V_k) \, dt - g_L (V - V_L) \, dt + dI_{\text{syn}}(V, t). \tag{3.1}$$

The synaptic input is given by $I_{\text{syn}}(V, t) = I_{\text{syn}}(t)$ or $I_{\text{syn}}(V, t) = aN^E(t) - aN^I(t)$ where $N^E(t)$ and $N^I(t)$ are Poisson processes with rate λ and $r\lambda$. The equations and parameters used in the HH model are standard (see [4, 31]).

We use the following set of parameters in simulations [4] for the IF-FHN model:

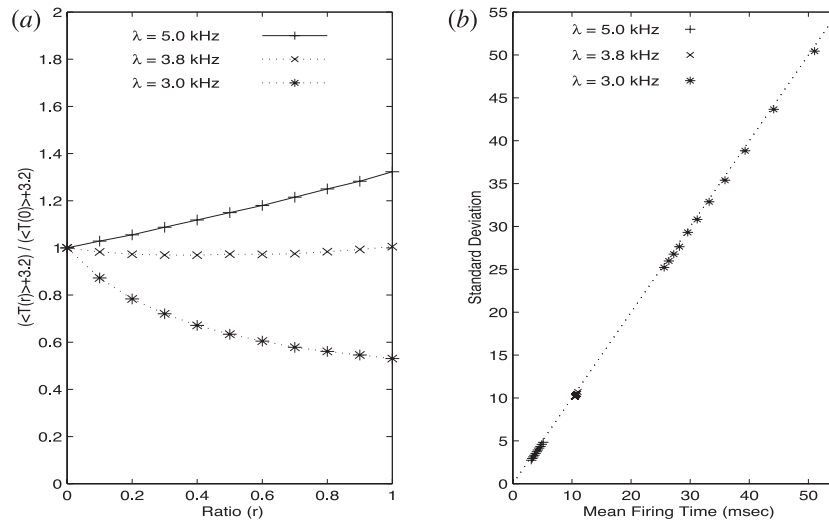$$\gamma = 100 \qquad \alpha = 0.2 \qquad \beta = 2.5$$

Figure 2. $a = 0.1$ for the IF-FHN model. (*a*) When $\lambda = 5.0$ kHz, $\langle T(0) \rangle + T_{\text{ref}} = 6.33$ ms; $\lambda = 3.8$ kHz, $\langle T(0) \rangle + T_{\text{ref}} = 14.36$ ms and $\lambda = 3.0$ kHz, $\langle T(0) \rangle + T_{\text{ref}} = 57.17$ ms. (*b*) Standard deviation of $T(r)$ versus $\langle T(r) \rangle$. It is easily seen that the standard deviation is almost equal to $\langle T(r) \rangle$, namely the efferent spike trains are Poisson process.

$v_{\text{thre}} = 1$ and $v_{\text{rest}} = 0$. Recently it has been shown in [16] that the HH (IF-FHN) model and the IF model behave in opposite ways when they receive correlated inputs.

In figure 2 we see that when the excitatory input frequency is high ($\lambda = 5$ kHz), the output firing rate is a decreasing function of inhibitory input rate, in agreement with our intuition. For example, when $r = 0$, $\langle T(0) \rangle + T_{\text{ref}} = 6.33$ ms and $r = 1$, $\langle T(1) \rangle + T_{\text{ref}} = 8.32$ ms. When the excitatory input frequency is around ($\lambda = 3.8$ kHz), the output firing rate is almost a constant function of inhibitory input rate. For example, when $r = 0$, $\langle T(0) \rangle + T_{\text{ref}} = 14.36$ ms and $r = 1$, $\langle T(1) \rangle + T_{\text{ref}} = 14.26$ ms. Further reducing the excitatory input rate shows the IBF phenomenon: *increasing inhibitory inputs increases neuronal firing rates*. For example, when $r = 0$, $\langle T(0) \rangle + T_{\text{ref}} = 57.17$ and $r = 1$, $\langle T(1) \rangle + T_{\text{ref}} = 29.87$.

Let us define a *critical input frequency* λ_c as the quantity which satisfies

$$\langle T(0) \rangle = \langle T(1) \rangle. \tag{3.2}$$

The numerical results of figure 2 tell us that $\lambda_c \sim 3.8$ kHz when $a = 0.1$.

It has been widely reported in the literature that increasing inhibitory input to a neuron could increase the variability of its output [3,4,8,9,27,28]. Figure 2(*a*) shows standard deviation versus mean firing time $\langle T(r) \rangle$. As we have reported before [4], the standard deviation of $T(r)$ almost equals its mean. Therefore in the no-IBF parameter regions, increasing inhibitory input induces an increase of its variability of output. However, in the parameter regions in which the IBF occurs, we see that now the CV of $T(r)$ is a decreasing function of r rather than an increasing function of r, in contrast to conventional theory in the literature. Note that

$$\text{CV} = \frac{\text{standard deviation}}{\text{mean + refractory period}}$$

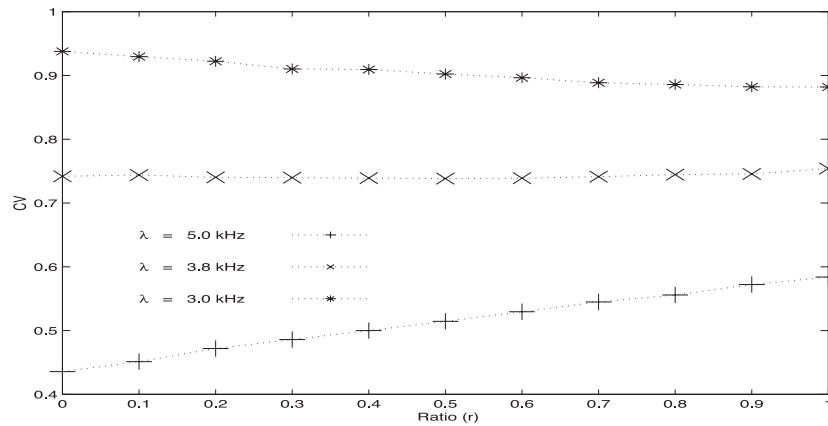where the refractory period is 8.32 ms. Hence figures 2 and 3 are consistent.

Figure 3. CV of IF-FHN model. CV could be either an increasing or a decreasing function of r.

4. Theoretical results

The main purpose of this section is to characterize the parameter regions of (a, λ) in which increasing inhibitory input increases neuronal firing rates. It thus gives us a complete picture of the IF model and IF-FHN model behaviour, and provides us with the answers to the 'when' and 'why' questions.

4.1. The IF model

As in [10], we could use the large-deviation theory [2] to estimate the mean firing time of the model. The obtained results are quite clear-cut. However, it is an approximation result and the accuracy is quite poor (not shown). Recently we have developed a rigorous theory [13] on the calculation of the mean firing time of the one-dimensional neuron model. We therefore adopt the rigorous approach here.

To this end, we first introduce some general notation. Consider a diffusion process defined by

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dB_t. \tag{4.1}$$

Let us introduce the following quantities:

$$
\begin{aligned}
s(x) &= \exp\left(-\int_0^x \frac{2\mu(y)}{\sigma^2(y)}\,dy \right) \\
m(x) &= \frac{1}{s(x)\sigma^2(x)} = \frac{\exp\left(\int_0^x \frac{2\mu(z)}{\sigma^2(z)}\,dz \right)}{\sigma^2(x)}
\end{aligned}
\tag{4.2}
$$

where m is the speed density and s is the scale function. We call a diffusion process positive-recurrent if $\int_{-\infty}^{\infty} m(x)\,dx < \infty$, which is equivalent to $\langle T \rangle < \infty$, where T is the first exit time of $(-\infty, V_{\text{thre}}]$. For a positive-recurrent process [23], its stationary distribution density is given by $\pi(x) \propto m(x)$.

The following conclusion is proved in [13].

Theorem 1. *For a positive-recurrent diffusion process X_t we have*

$$\langle T \rangle = 2 \int_{V_{\text{rest}}}^{V_{\text{thre}}} s(u)\,du \cdot \int_{-\infty}^{V_{\text{rest}}} m(u)\,du + 2 \int_{V_{\text{rest}}}^{V_{\text{thre}}} \left(\int_y^{V_{\text{thre}}} s(u)\,du \right) \cdot m(y)\,dy$$

$$= 2 \int_{V_{\text{rest}}}^{V_{\text{thre}}} \left(\int_{-\infty}^{y} m(u) \, du \right) \cdot s(y) \, dy. \tag{4.3}$$

According to the definition of the scale function we have

$$s(x) = \exp \left(\frac{Lx^2 - 2a\lambda(1-r)x}{a^2\lambda(1+r)} \right) \tag{4.4}$$

and therefore

$$\langle T(r) \rangle = 2 \int_{V_{\text{rest}}}^{V_{\text{thre}}} \frac{1}{a^2\lambda(1+r)} \exp \left(\frac{Lx^2 - 2a\lambda x(1-r)}{a^2\lambda(1+r)} \right)$$

$$\times \left(\int_{-\infty}^{x} \exp \left(-\frac{Lu^2 - 2a\lambda u(1-r)}{a^2\lambda(1+r)} \right) du \right) dx$$

$$= \frac{2}{L} \int_{\frac{V_{\text{rest}}\sqrt{L}}{a\sqrt{\lambda(1+r)}} - \frac{\sqrt{\lambda}(1-r)}{\sqrt{L(1+r)}}}^{\frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{\lambda(1+r)}} - \frac{\sqrt{\lambda}(1-r)}{\sqrt{L(1+r)}}} \left[\exp(x^2) \int_{-\infty}^{x} \exp(-u^2) \, du \right] dx. \tag{4.5}$$

For simplicity of notation let us assume that $V_{\text{rest}} = 0$. We have

$$\langle T(0) \rangle = \frac{2}{L} \int_{-\frac{\sqrt{\lambda}}{\sqrt{L}}}^{\frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{\lambda}} - \frac{\sqrt{\lambda}}{\sqrt{L}}} \left[\exp(x^2) \int_{-\infty}^{x} \exp(-u^2) \, du \right] dx$$

$$\langle T(1) \rangle = \frac{2}{L} \int_{0}^{\frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{2\lambda}}} \left[\exp(x^2) \int_{-\infty}^{x} \exp(-u^2) \, du \right] dx. \tag{4.6}$$

Taking the limit $\lambda \to 0$ we obtain

$$\langle T(0) \rangle \sim \frac{2}{L} \int_{0}^{\frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{\lambda}}} \left[\exp(x^2) \int_{-\infty}^{x} \exp(-u^2) \, du \right] dx$$

$$> \langle T(1) \rangle = \frac{2}{L} \int_{0}^{\frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{2\lambda}}} \left[\exp(x^2) \int_{-\infty}^{x} \exp(-u^2) \, du \right] dx \tag{4.7}$$

since the integral interval for $\langle T(1) \rangle$ is

$$\left[0, \frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{2\lambda}} \right] \subset \left[0, \frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{\lambda}} \right]$$

which is the integral interval for $\langle T(0) \rangle$. It is easily seen that there is a $\lambda > 0$ such that

$$\langle T(0) \rangle < \langle T(1) \rangle. \tag{4.8}$$

Furthermore, all the functions

$$\frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{2\lambda}} \qquad \frac{V_{\text{thre}}\sqrt{L}}{a\sqrt{\lambda}} \qquad \frac{\sqrt{\lambda}}{\sqrt{L}}$$

and therefore the integral bounds, are monotonic functions of λ and thus both $\langle T(0) \rangle$ and $\langle T(1) \rangle$ are monotonic functions of λ. As shown in appendix A, $\langle T(0) \rangle$ and $\langle T(1) \rangle$ are convex functions of λ. Hence we have the following theorem.

Theorem 2. *There is a unique critical input frequency $0 < \lambda_c < \infty$ satisfying*

$$\langle T(0) \rangle = \langle T(1) \rangle. \tag{4.9}$$

The IF model has been studied for over a century [22]. Nevertheless it seems that the result given by equation (4.5) is the simplest rigorous result in the literature [25, 31], and could be easily employed to explore some related issues [21, 22]. In the next section, we apply theorem 1 to the IF-FHN model. In fact, theorem 1 could be employed to tackle any one-dimensional neuronal model: for example, that of the θ-neuron [6].

4.2. IF-FHN model

Applying theorem 1 to the IF-FHN model, we obtain

$$
s(x) = \exp\left[\int_0^x \frac{2L(y)y - 2\mu}{\sigma^2}\,\mathrm{d}y\right]
$$

$$
= \exp\left[\frac{\gamma}{2\sigma^2}x^4 - \frac{2(V_{\text{thre}} + \alpha)\gamma}{3\sigma^2}x^3 + \frac{\alpha\gamma\beta V_{\text{thre}} + 1}{\sigma^2\beta}x^2 - \frac{2\mu x}{\sigma^2}\right] \qquad (4.10)
$$

and

$$
m(x) = \frac{1}{\sigma^2}\exp\left[-\frac{\gamma}{2\sigma^2}x^4 + \frac{2(V_{\text{thre}} + \alpha)\gamma}{3\sigma^2}x^3 - \frac{\alpha\gamma\beta V_{\text{thre}} + 1}{\sigma^2\beta}x^2 + \frac{2\mu x}{\sigma^2}\right]. \qquad (4.11)
$$

However, a fully theoretical treatment of these quantities, as in the previous section, is difficult because of the nonlinearity of the leakage coefficient.

From the data shown in figure 2 we might envisage that Kramer's formula, a special case of the large-deviation theory (see [2] and references therein for details) can predict the model behaviour. Kramer's model [26] reads

$$
\langle T(r)\rangle \sim \frac{2\pi}{\sqrt{H''(v_{\min})|H''(v_{\max})|}}\exp\left[2(H(v_{\max}) - H(v_{\min}))/\sigma^2\right] \qquad (4.12)
$$

where

$$
H(v) = \gamma\left[\frac{1}{4}v^4 - (\alpha + 1)\frac{1}{3}v^3 + \frac{1}{2}\alpha v^2\right] + \frac{1}{2\beta}v^2 - \mu v \qquad (4.13)
$$

and $v_{\max}$, $v_{\min}$ are the local maximum and left local minimum of the potential well H. Hence $v_{\max}$ is the exact threshold of the IF-FHN model. Different from the IF model, now Kramer's formula gives quite a good approximation result, as shown in the next section. Due to the computational advantage of Kramer's formula, we concentrate on it in the following theoretical development for the IF-FHN model.

Another interesting feature of the IF-FHN model, different from the IF model, is that the exact threshold changes in response to synaptic inputs. Here α, β and γ take the values given before. From figure 4 we see that, to ensure a two-well system, we should have $v_{\max} \geqslant 0.096\,685$. When $\mu \geqslant 0.940\,9837$, the model becomes a dynamical system with a single well, i.e. the stimulus is a supra-threshold one.

Now we use a geometrical method to prove, under certain circumstances, that for any given $a > 0$, there is a $\lambda_{\text{c}} > 0$ such that

$$
\langle T(0)\rangle = \langle T(1)\rangle. \qquad (4.14)
$$

If we further know the uniqueness of $\lambda_{\text{c}} > 0$ satisfying equation (4.14), we can assert that when $\lambda < \lambda_{\text{c}}$, $\langle T(0)\rangle > \langle T(1)\rangle$ and when $\lambda > \lambda_{\text{c}}$, $\langle T(0)\rangle < \langle T(1)\rangle$.

Consider figure 4: the curve is the function $H'(v) + \mu$ which is thus independent of synaptic inputs. Therefore, for any given $\mu \geqslant 0$, intersections with the horizontal line $y = \mu$ give us $v_{\min}$ and $v_{\max}$ of the potential well with input μ, if $v_{\min}$ and $v_{\max}$ exist, i.e. when the stimulus is a subthreshold one. In terms of this scenario, we could calculate each term in Kramer's formula.

Let us first consider the function

$$
F(\mu) = \sqrt{H''(v_{\min})|H''(v_{\max})|}
$$

which depends only on μ. Figure 5(a) depicts $F(\mu)$ versus μ. It is readily seen that $F(\mu)$ is a decreasing function of μ.
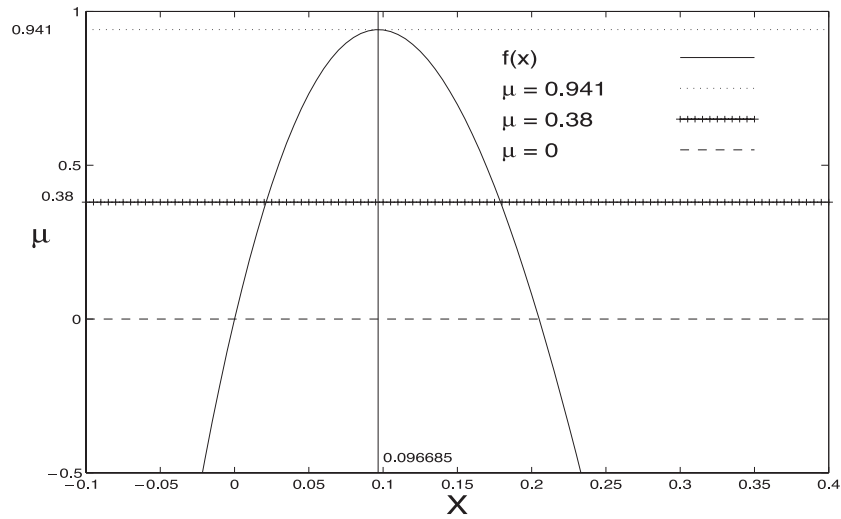
Figure 4. An explanation of the geometrical method for the IF-FHN model. $f(x) = \gamma(x^3 - (v_{\text{thre}} + \alpha)x^2 + \alpha v_{\text{thre}}x) + x/\beta$. When $a = 0.1$ we know from the previous section that if $0 < \mu < 0.38$ (indicated by points), IBF occurs.
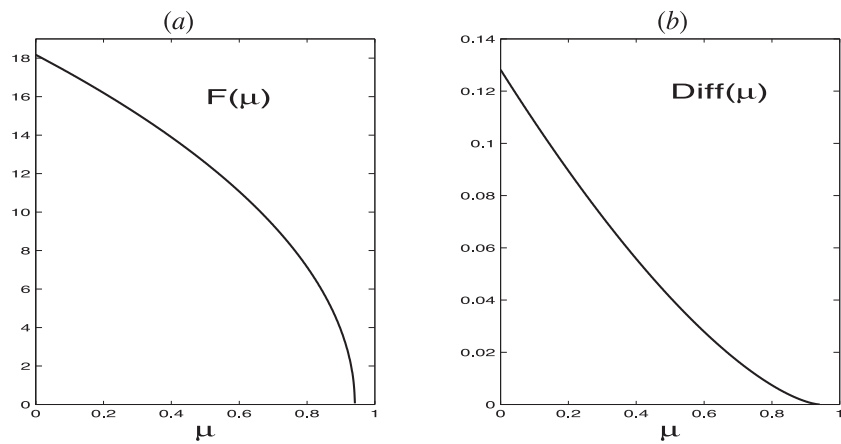

Figure 5. IF-FHN model. (*a*) F versus μ. $F(\mu)$ is a decreasing function for $\mu \in [0, 0.940\,987]$. (*b*) $\text{Diff}(\mu) = H(v_{\text{max}}) - H(v_{\text{min}})$ versus μ. Note that Diff depends only on μ as well.

Now consider the second term (figure 5(*b*))

$$\exp\left[\frac{2(H(v_{\text{max}}) - H(v_{\text{min}}))}{\sigma^2}\right] = \exp\left[\frac{2(H(v_{\text{max}}) - H(v_{\text{min}}))}{a^2\lambda(1+r)}\right].$$

When $r = 0$ the term above becomes

$$\exp\left[\frac{2(H(v_{\text{max}}) - H(v_{\text{min}}) + a\lambda(v_{\text{max}} - v_{\text{min}}) - a\lambda(v_{\text{max}} - v_{\text{min}}))}{a^2\lambda}\right]$$

$$= \exp\left[\frac{2(H(v_{\text{max}}) - H(v_{\text{min}}) + a\lambda(v_{\text{max}} - v_{\text{min}}))}{a^2\lambda}\right]\exp\left[-\frac{v_{\text{max}} - v_{\text{min}}}{a}\right].$$

$$(4.15)$$

When $r = 1$ the term is

$$\exp\left[\frac{2(H(v_{\max}) - H(v_{\min}))}{2a^2\lambda}\right]. \tag{4.16}$$

It is easily seen that when λ tends to zero, $H(v_{\max}) - H(v_{\min}) + a\lambda(v_{\max} - v_{\min}), r = 0$ is identical to $H(v_{\max}) - H(v_{\min}), r = 1$. Hence the first term of equations (4.15), (4.16) tends to infinity as $\lambda \to 0$. The second term in equation (4.15) is bounded for any λ. Combining with the behaviour of $F(\mu)$, which is a decreasing function of μ, we conclude that

$$\langle T(0)\rangle > \langle T(1)\rangle$$

provided that λ is small.

The proof of

$$\langle T(0)\rangle < \langle T(1)\rangle$$

for large λ is trivial. We know from theorem 1 that when λ is large enough

$$\langle T(0)\rangle = o(\langle T(1)\rangle).$$

A more delicate issue is the uniqueness of λ_c which we have not proved, though our numerical simulations suggest the validity of the claim.

Theorem 3. *For the IF-FHN model, suppose that $F(\mu)$ is a decreasing function, then there is a λ_c with the property that*

$$\langle T(0)\rangle = \langle T(1)\rangle.$$

Finally, we point out that although here we confine ourselves to IF-FHN model and therefore the potential function H is given by equation (4.13), all our proofs do not depend on the concrete form of H.

5. Numerical results

For the IF model, equation (4.5) agrees perfectly with numerical results, as shown in figure 6(*a*). Parameters used in the simulations are $\lambda = 10$ kHz, $a = 0.5$ mV, $V_{\text{thre}} = 20$ mV, $V_{\text{rest}} = 0$ and $L = 1/20.2$.

Figure 6(*b*) shows a comparison between numerical simulations and results obtained in terms of Kramer's formula (see also figures in [10]). We can see that there is a discrepancy between them, although in general they agree with each other. Moreover, numerical results are in agreement with large-deviation theory [2] which indicates that

$$\lim_{\lambda \to 0} \sigma^2 \log\langle T(r)\rangle = 2[H(v_{\max}) - H(v_{\min})].$$

Let us now, in terms of Kramer's formula, characterize the whole parameter regions of (a, λ) in which inhibitory inputs increase neuronal firing rates for the IF-FHN model.

Figure 7 shows the curve of λ_c versus a for $0.02 \leqslant a \leqslant 0.1$. For example, when $a = 0.08$ we have $\lambda_c = 4.6$ kHz and the neuron fires at a rate of 38 Hz. Namely when $\lambda \leqslant \lambda_c$, increasing inhibitory input enhances the neuronal firing rates. In general, we see that when the input frequency is small, the inhibitory input increases the neuronal activity. Therefore, with the help of inhibitory inputs, a neuron actively modulates its activity: when the input is high, inhibitory input decreases its firing; when input is low, inhibitory input increases its firing.
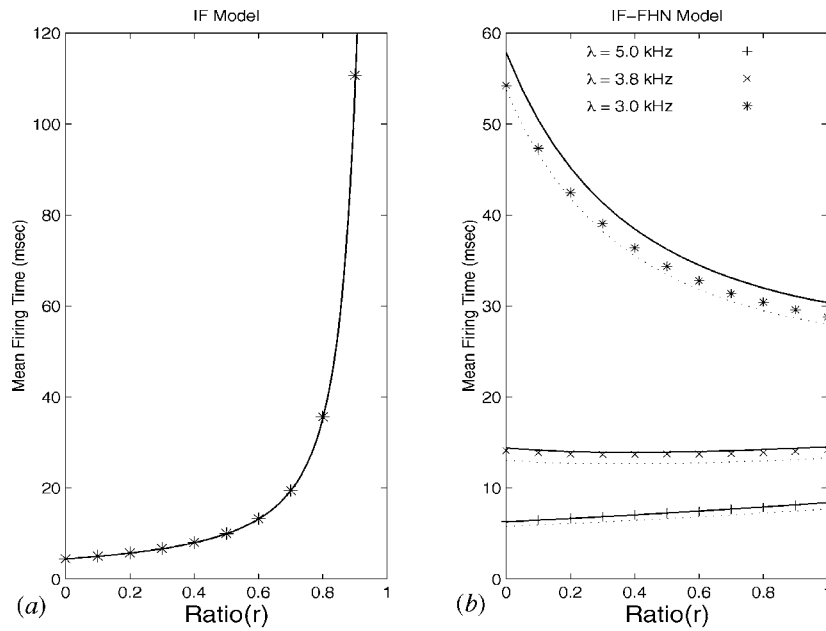
Figure 6. (*a*) A comparison between equation (4.5) and numerical results of the IF model. Points are simulation results; the curve represents the exact values by theory. (*b*) A comparison between Kramer's formula and numerical results of the IF-FHN model. Points are simulation results; dotted curves represent the approximations by Kramer's formula; solid curves represent the direct numerical evaluation of the integral in theorem 1.

6. Correlated inputs

In previous sections we considered the IF model and IF-FHN model with independent inputs. Nevertheless, as we have pointed out before [11], a neuron usually receives spatially correlated inputs, rather than independent inputs [29, 33]. In this section we consider the impact of correlated inputs on the IBF phenomenon.

We briefly review the formulation of correlated inputs. Suppose that a neuron is subjected to N_E excitatory synaptic inputs and N_I inhibitory synaptic input. Let $\bar{\lambda}$ be the firing rate of each excitatory synapse and ρ be the correlation coefficient between excitatory (inhibitory) synapses. For simplicity of notation we assume that $N_E = N_I$. Then we have

$$
\begin{aligned}
\mu &= a\bar{\lambda}N_E(1 - r) \\
\sigma^2 &= a^2\bar{\lambda}N_E(1 + \rho(N_E - 1))(1 + r).
\end{aligned}
\tag{6.1}
$$

Hence $\lambda = \bar{\lambda}N_E$ and when $\rho = 0$, the neuron receives independent inputs. The correlation between synapses increases the noise term, but keeps the mean input unchanged. As we have pointed out before, the IBF phenomenon is due to a competition between the noise force and the deterministic force. We can naturally expect that increasing correlation, or the noise force, will facilitate the IBF phenomenon.

Figure 8 presents numerical results further explaining the conclusions above. Even with a small correlation ($\rho = 0.05$), the IBF region for the IF-FHN model is considerably enlarged, compared with the case of $\rho = 0.0$. Assume that a neuron fires with its maximum rate of 500 Hz: we see from figure 8 that when $a \geqslant 0.03$ (approximately), $\lambda < \lambda_c$, namely all physiologically plausible parameters of inputs are inside the IBF region.
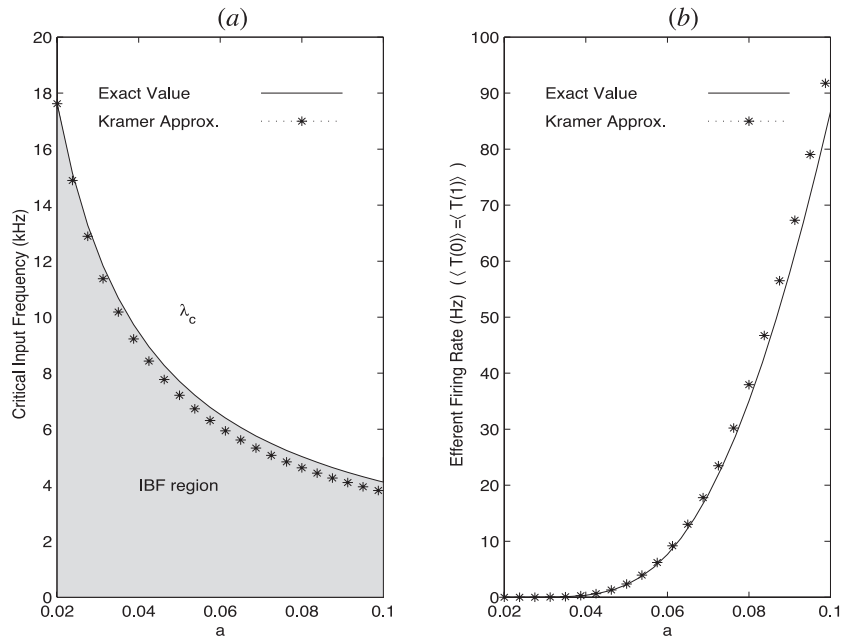
Figure 7. IF-FHN model. (*a*) λ_c versus *a* for the IF-FHN model. (*b*) Efferent frequency when $\lambda = \lambda_c$ versus *a*.

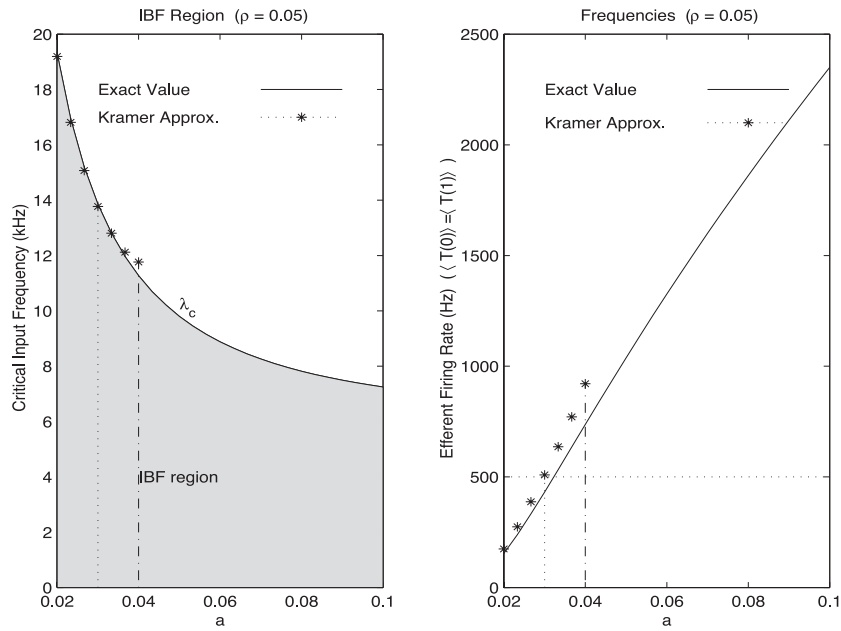**Figure 8.** λ_c versus *a* for the IF-FHN model and efferent frequencies $\langle T(0) \rangle = \langle T(1) \rangle$ when $\lambda = \lambda_c$ versus *a*. Note that to ensure the efferent frequencies are in physiologically plausible ranges, i.e. efferent firing rate $\leqslant 500$ Hz, when $a > 0.03$ ($\lambda_c < 13.856$), λ is in the IBF region.

As we can see from figure 8 and previous numerical results, in the reasonable parameter ranges, the Kramer formula offers a fairly good approximation. Denoting $\lambda_c^*(\rho)$ as the Kramer approximation of the critical input frequency at which

$$\langle T(0) \rangle = \langle T(1) \rangle$$

we have the following theorem.

Theorem 4. *For the IF-FHN model, with the assumption of unique critical input frequency, if $\rho_1 > \rho_2$, then*

$$\lambda_c^*(\rho_1) > \lambda_c^*(\rho_2).$$

The theorem above follows from the simple observation that $F(\mu)$ is independent of ρ and

$$\exp\left[\frac{2(H(v_{\max}) - H(v_{\min}))}{a^2 \bar{\lambda} N_E (1 + \rho_1(N_E - 1))(1 + r)}\right] < \exp\left[\frac{2(H(v_{\max}) - H(v_{\min}))}{a^2 \bar{\lambda} N_E (1 + \rho_2(N_E - 1))(1 + r)}\right]$$

when $\rho_1 > \rho_2$.

Similar conclusions hold true for the IF model as stated in the following theorem.

Theorem 5. *For the IF model, if $\rho_1 > \rho_2$ then*

$$\lambda_c(\rho_1) > \lambda_c(\rho_2).$$

The proof is postponed to appendix B.

7. Discussion

We answer the following questions in this paper: why and when does increasing inhibitory inputs increase neuronal firing rates? For the IF model, we show that there is a unique input frequency λ_c at which the efferent firing rate of a neuron is identical when the cell receives purely excitatory inputs or exactly balanced inhibitory and excitatory inputs. For the IF-FHN model, by Kramer's formula, we prove that when input frequency is low enough the model increases its efferent firing rate when inhibitory inputs are added to the model. Our results provide a theoretical foundation for the IBF phenomenon and might alter our traditional views on stochastic inputs of a neuron.

We point out that the mechanism of IBF described here is totally different from that in [30]. In [30] the authors consider a network of excitatory and inhibitory neurons and find that increasing the direct external inhibitory input to the inhibitory interneurons, without directly affecting any other part of the network, can, in some circumstances, cause the interneurons to increase their firing rates. This is essentially a network phenomenon, but the IBF phenomenon, as we have emphasized before, is observable for single neuron, whether it is an inhibitory (interneuron) or an excitatory neuron. Furthermore, the IBF phenomenon cannot be observed if the input is deterministic which is the case in [30]. The IBF is due to the ergodic property of a system driven by noise.

The classical approach to neural networks is that excitatory neurons are the main information carriers. It is thought that neurons receive virtually all significant input from excitatory inputs, and that inhibitory neurons interact only locally and responsively. The stronger the (excitatory) inputs, the higher the output rate. When modelling, the conventional view is that information is received solely through excitatory inputs (EPSPs) [1], and that neurons process this information either by integration or coincidence detection. Inhibitory inputs are often thought of as broad dampening influences usually reducing neuronal

responsiveness to all input. The role of inhibitory inputs has been extensively discussed in recent years, and various functions have been postulated. These include the production of a more graded response to changes in input, and facilitating neuronal synchronization [32]. Substantial inhibitory input has also been demonstrated to be one factor which can induce increased random variability of neuronal output [28]. Here we find a more surprising effect of inhibitory input: actually increasing the neuronal firing rate. In [4], we reported that increasing inhibitory inputs had almost no impact on the output firing rate and CF of output interspike interval for the HH and the FitzHugh–Nagumo models. Results in this paper add an important new dimension to these results by showing that increasing inhibitory inputs can even enhance neuronal activity. The IBF phenomenon presented here provides another surprising feature of neuronal models, and alters our view of how information might be processed by neurons. It leads us to predict the existence of mechanisms exploring IBF to increase neuronal responsiveness and sensitivity, involving the inhibitory circuits widely present in the brain.

Acknowledgments

Appendix A. The Convexity of the mean firing time as a function of λ

By examining the integral form of the mean firing time in equation (4.6), it is quite clear that we need only to show the convexity of the following function:

$$f(y) = \int_{-\frac{1}{y}}^{cy-\frac{1}{y}} g(x)\, dx \qquad (y > 0) \tag{A.1}$$

where $c > 0$, and

$$g(x) = e^{x^2} \int_{-\infty}^{x} e^{-u^2}\, du. \tag{A.2}$$

Now

$$g'(x) = 2xg(x) + 1 \tag{A.3}$$

which is positive for $x \geqslant 0$. With integration by part, it can be easily shown that

$$g'(x) = -xe^{x^2} \int_{-\infty}^{x} \frac{1}{u^2} e^{-u^2}\, du \qquad (x < 0). \tag{A.4}$$

Hence $g(x)$ is strictly increasing for $x \in R$, and

$$g(x) = -\frac{1}{2x}\left(1 + xe^{x^2} \int_{-\infty}^{x} \frac{1}{u^2} e^{-u^2}\, du\right) \qquad (x < 0). \tag{A.5}$$

Furthermore, let

$$h(x) = xg(x). \tag{A.6}$$

Obviously $h(x)$ is an increasing function for $x \geqslant 0$; and when $x < 0$,

$$
\begin{aligned}
h'(x) &= g(x) + 2x^2 g(x) + x \\
&= g(x) - 2x^2 \frac{1}{2x}\left(1 + xe^{x^2}\int_{-\infty}^{x}\frac{1}{u^2}e^{-u^2}\,du\right) + x \\
&= g(x) - x^2 e^{x^2}\int_{-\infty}^{x}\frac{1}{u^2}e^{-u^2}\,du \\
&= e^{x^2}\int_{-\infty}^{x}\left(1 - \frac{x^2}{u^2}\right)e^{u^2}\,du > 0 \qquad (x < 0).
\end{aligned}
$$

So $xg(x)$ is also a strictly increasing function for $x < 0$. Thus

$$
f'(y) = g\left(cy - \frac{1}{y}\right)\left(c + \frac{1}{y^2}\right) - \frac{1}{y^2}g\left(-\frac{1}{y}\right) \tag{A.7}
$$

$$
\begin{aligned}
f''(y) &= g'\left(cy - \frac{1}{y}\right)\left(c + \frac{1}{y^2}\right)^2 + g\left(cy - \frac{1}{y}\right)\left(-\frac{2}{y^3}\right) \\
&\quad + \frac{2}{y^3}g\left(-\frac{1}{y}\right) - \frac{1}{y^4}g'\left(-\frac{1}{y}\right) \tag{A.8}
\end{aligned}
$$

$$
\begin{aligned}
&= 2g\left(cy - \frac{1}{y}\right)\left[\left(cy - \frac{1}{y}\right)\left(c + \frac{1}{y^2}\right)^2 - \frac{1}{y^3}\right] \\
&\quad + \frac{2}{y^3}g\left(-\frac{1}{y}\right)\left[1 + \frac{1}{y^2}\right] + \left[c^2 + \frac{2c}{y^2}\right] \qquad \text{by (A.3)}. \tag{A.9}
\end{aligned}
$$

By the mean value theorem

$$
g\left(-\frac{1}{y}\right) - g\left(cy - \frac{1}{y}\right) = -g'(\zeta)cy \qquad \zeta \in \left[-\frac{1}{y}, -\frac{1}{y} + cy\right] = -[2\zeta g(\zeta) + 1]cy. \tag{A.10}
$$

Hence

$$
\begin{aligned}
f''(y) &= 2g\left(cy - \frac{1}{y}\right)\left(cy - \frac{1}{y}\right)\left(c + \frac{1}{y^2}\right)^2 - \frac{2}{y^3}[2\zeta g(\zeta) + 1]cy \\
&\quad + \frac{2}{y^5}g\left(-\frac{1}{y}\right) + \left(c^2 + \frac{2c}{y^2}\right) \\
&= 2g\left(cy - \frac{1}{y}\right)\left(cy - \frac{1}{y}\right)\left(c + \frac{1}{y^2}\right)^2 - \frac{4\zeta g(\zeta)c}{y^2} + \frac{2}{y^5}g\left(-\frac{1}{y}\right) \\
&= 2\left(c + \frac{1}{y^2}\right)^2\left[\left(cy - \frac{1}{y}\right)g\left(cy - \frac{1}{y}\right) - \zeta g(\zeta)\right] \\
&\quad + \frac{2}{y^4}\left[\zeta g(\zeta) - \left(-\frac{1}{y}\right)g\left(-\frac{1}{y}\right)\right] + c^2(1 + 2\zeta g(\zeta)) \\
&> 0 \qquad \text{since } g(x) \text{ and } xg(x) \text{ are increasing functions and (A.3)}. \tag{A.11}
\end{aligned}
$$

This ensures the convexity of $f(x)$.

Appendix B. Proof of theorem 5

Proof. For any given $1 > \rho_1 > \rho_2 > 0$, let $\theta_i = (1 + \rho_i(N_E + 1))$, $i = 1, 2$, and $\lambda = \bar{\lambda}N_E$. Then we have $\theta_1 > \theta_2 > 1$.

For given ρ_i, $i = 1, 2$, and in terms of equation (4.6), we have

$$\langle T(0) \rangle_i = \int_{-\frac{\sqrt{\lambda/\theta_i^2}}{\sqrt{L}}}^{\frac{p}{\sqrt{\lambda\theta_i^2}} - \frac{\sqrt{\lambda/\theta_i^2}}{\sqrt{L}}} g(x)\,\mathrm{d}x \overset{\text{def}}{=} G_0(\lambda, \theta_i)$$

(B.1)

$$\langle T(1) \rangle_i = \int_0^{\frac{p}{\sqrt{2\lambda\theta_i^2}}} g(x)\,\mathrm{d}x \overset{\text{def}}{=} G_1(\lambda, \theta_i) \qquad i = 1, 2$$

where $p = V_{\text{thre}}\sqrt{L}/a$, and

$$g(x) = \frac{2}{L}\exp(x^2)\int_{-\infty}^x \exp(-u^2)$$

is a positive increasing function.

Denote $\lambda_i = \lambda_c(\rho_i)$ as the unique critical input frequency corresponding to ρ_i, $i = 1, 2$. So

$$\langle T(0) \rangle_i > \langle T(1) \rangle_i \qquad \lambda < \lambda_i \qquad i = 1, 2.$$

(B.2)

From the proof of theorem 2, we know that $G_0(\lambda, \theta_i) = \langle T(0) \rangle_i$ and $G_1(\lambda, \theta_i) = \langle T(1) \rangle_i$, $i = 1, 2$, are increasing convex functions of λ. Therefore, to prove $\lambda_1 > \lambda_2$, it is enough to show that

$$\langle T(0) \rangle_1 > \langle T(1) \rangle_1 \qquad \text{when} \quad \lambda = \lambda_2$$

(B.3)

i.e.

$$G_0(\lambda_2, \theta_1) > G_1(\lambda_2, \theta_1).$$

(B.4)

But

$$G_0(\lambda_2, \theta_1) = \int_{-\frac{\sqrt{\lambda_2/\theta_1^2}}{\sqrt{L}}}^{\frac{p}{\sqrt{\lambda_2\theta_1^2}} - \frac{\sqrt{\lambda_2/\theta_1^2}}{\sqrt{L}}} g(x)\,\mathrm{d}x$$

(B.5)

$$> \int_{-\frac{\sqrt{\lambda_2\theta_1^2}}{\sqrt{L}}}^{\frac{p}{\sqrt{\lambda_2\theta_1^2}} - \frac{\sqrt{\lambda_2\theta_1^2}}{\sqrt{L}}} g(x)\,\mathrm{d}x$$

(B.6)

$$= G_0(\lambda_2\theta_1^2, 1)$$

(B.7)

$$= G_1(\lambda_2\theta_1^2, 1) = G_1(\lambda^*, \theta_2) \qquad \lambda^* = \lambda_2\theta_1^2\theta_2^2 > \lambda_2$$

(B.8)

$$> G_1(\lambda_2, \theta_2) = G_0(\lambda_2, \theta_2)$$

(B.9)

$$> G_1(\lambda_2, \theta_1)$$

(B.10)

where inequality (B.6) is because the integrant $g(x)$ is strictly increasing, and the integral range is shifted towards negative infinity; inequality (B.10) is from the fact that $\theta_1 > \theta_2$ and the integral form of $G_1(\lambda, \theta_i)$ in (B.1); equality in equations (B.7), (B.8) and (B.9) is due to the assumption in the theorem that λ_2 is the critical point.

This completes our proof. □

References

[1] Abbott L F, Varela J A, Sen K and Nelson S B 1997 Synaptic depression and cortical gain control *Science* **275** 220–3
[2] Albeverio S, Feng J and Qian M 1995 Role of noise in neural networks *Phys. Rev.* E **52** 6593–606

[3] Brown D and Feng J 1999 Is there a problem matching model and real CV(ISI)? *Neurocomputing* **26–27** 117–22
[4] Brown D, Feng J and Feerick S 1999 Variability of firing of Hodgkin–Huxley and FitzHugh–Nagumo neurons with stochastic synaptic input *Phys. Rev. Lett.* **82** 4731–4
[5] Collins J J, Chow C C and Imhoff T T 1995 Stochastic resonance without tuning *Nature* **376** 236–8
[6] Ermemtrout B 1996 Type I membranes, phase resetting curves, and synchrony *Neural Comput.* **8** 979–1001
[7] Feng J 1997 Behaviours of spike output jitter in the integrate-and-fire model *Phys. Rev. Lett.* **79** 4505–8
[8] Feng J and Brown D 1998 Impact of temporal variation and the balance between excitation and inhibition on the output of the perfect integrate-and-fire model *Biol. Cybern.* **78** 369–76
[9] Feng J and Brown D 1999 Coefficient or variation greater than 0.5 how and when? *Biol. Cybern.* **80** 291–7
[10] Feng J and Brown D 2000 Integrate-and-fire models with nonlinear leakage *Bull. Math. Biol.* **62** 467–81
[11] Feng J and Brown D 2000 Impact of correlated inputs on the output of the integrate-and-fire model *Neural Comput.* **12** 671–92
[12] Feng J, Brown D and Li G 2000 Synchronization due to common pulsed input in Stein's model *Phys. Rev.* E **61** 2987–95
[13] Feng J and Li G 2001 Impact of geometrical structures on the output of neuronal models—from spikes to bursts *Neural Comput.* **13** at press
[14] Feng J and Tirozzi B 2000 Stochastic resonance tuned by correlations in neuronal models *Phys. Rev.* E **61** 4207–11
[15] Feng J and Cassia-Moura R 1999 Output of a neuronal population code *Phys. Rev.* E **59** 7246–9
[16] Feng J and Zhang P 2001 The behaviour of integrate-and-fire and Hodgkin–Huxley models with correlated inputs *Phys. Rev.* E **63** 051902
[17] FitzHugh R 1961 Impulses and physiological states in theoretical models of nerve membrane *Biophys. J.* **1** 445–66
[18] Gerstner W, Kreiter A K, Markram H and Herz A V M 1997 Neural codes: firing rates and beyond *Proc. Natl Acad. Sci. USA* **94** 12740–1
[19] Gammaitoni L, Hänggi P, Jung P and Marchesoni F 1998 Stochastic resonance *Rev. Mod. Phys.* **70** 224–87
[20] Hodgkin A L and Huxley A F 1952 Currents carried by sodium and potassuim ions through the membrane of hte giant axon of Loligo *J. Physiol.* **116** 449–506
[21] Holt G R and Koch C 1997 Shunting inhibition does not have a divisive effect on firing rates *Neural Comput.* **9** 1001–13
[22] Koch C 1999 *Biophysics of Computation* (Oxford: Oxford University Press)
[23] Karlin S and Taylor H M 1982 *A Second Course in Stochastic Processes* (New York: Academic)
[24] Nagumo J S, Arimoto S and Yoshizawa S 1962 An active pulse transmission line simulating nerve axon *Proc. IRE* **50** 2061–70
[25] Ricciardi L M and Sato S 1990 Diffusion process and first-passage-times problems *Lectures in Applied Mathematics and Informatics* ed L M Ricciardi (Manchester: Manchester University Press)
[26] Risken S 1989 *The Fokker–Planck Equation* (Berlin: Springer)
[27] Softky W and Koch C 1993 The highly irregular firing of cortical-cells is inconsistent with temporal integration of random EPSPs *J. Neurosci.* **13** 334–50
[28] Shadlen M N and Newsome W T 1994 Noise, neural codes and cortical organization *Curr. Opin. Neurobiol.* **4** 569–79
[29] Stevens C F and Zador A M 1998 Input synchrony and the irregular firing of cortical neurons *Nat. Neurosci.* **1** 210–7
[30] Tsodyks M V, Skaggs W E, Sejnowski T J and McNaughton B L 1997 Paradoxical effects of external modulation of inhibitory interneurons *J. Neurosci.* **17** 4382–8
[31] Tuckwell H C 1988 *Introduction to Theoretical Neurobiology* vol 2 (Cambridge: Cambridge University Press)
[32] van Vereeswijk C, Abbott L F and Ermentrout G B 1994 When inhibition not excitation synchronizes neural firing *J. Comput. Neurosci.* **1** 313–21
[33] Zohary E, Shadlen M N and Newsome W T 1994 Correlated neuronal discharge rate and its implications for psychophysical performance *Nature* **370** 140–3