Accuracy of rate coding: When shorter time window and higher spontaneous activity help

Marie Levakova,[1,*] Massimiliano Tamborrino,[2] Lubomir Kostal,[1] and Petr Lansky[1]

[1]*Department of Computational Neuroscience,*
Institute of Physiology of the Czech Academy of Sciences
Videnska 1083, 14220 Prague 4, Czech Republic

[2]*Institute for Stochastics, Johannes Kepler University Linz*
Altenbergerstraße 69, 4040 Linz, Austria

(Dated: January 20, 2017)

Abstract

It is widely accepted that neuronal firing rates contain a significant amount of information about the stimulus intensity. Nevertheless, theoretical studies on the coding accuracy inferred from the exact spike counting distributions are rare. We present an analysis based on the number of observed spikes assuming the stochastic perfect integrate-and-fire model with a change point, representing the stimulus onset, for which we calculate the corresponding Fisher information to investigate the accuracy of rate coding. We analyze the effect of changing the duration of the time window and the influence of several parameters of the model, in particular the level of the presynaptic spontaneous activity and the level of random fluctuation of the membrane potential, which can be interpreted as noise of the system. The results show that the Fisher information is nonmonotone with respect to the length of the observation period. This counter-intuitive result is caused by the discrete nature of the count of spikes. We observe also that the signal can be enhanced by noise, since the Fisher information is nonmonotone with respect to the level of spontaneous activity and, in some cases, also with respect to the level of fluctuation of the membrane potential.

* marie.levakova@fgu.cas.cz

I. INTRODUCTION

The question how the information is encoded in the brain activity remains one of the major unresolved problems in neuroscience. A pioneering work in this direction was done by Adrian [1] who did experiments on the stretch receptor in a muscle spindle and demonstrated that the number of spikes emitted by the receptor neuron increases with the force applied to the muscle. Since then, the neuronal firing rate has been traditionally thought to contain relevant information about external stimuli. The firing rate is usually determined by counting the number of spikes that occur within a predefined time window [2]. The irregularities encountered in real spike trains make the determination of the firing rate complicated, and thus a sufficiently long time window is necessary to achieve a satisfactory precision.

It is clear that an approach based only on the firing rate neglects all the information possibly contained in the exact timing of the spikes [3–5], leading to the concept of temporal coding [6–10]. One of the common objections against reducing the complexity of neural coding only to firing rates is based on behavioral experiments, which suggest that reaction times are often rather short [11] and thus are against the common understanding that the reliability of rate coding increases proportionally to the applied time window.

In the last decades, the performance of both temporal and rate coding was investigated for various settings, ranging from single neurons to neuronal networks, see for example [12–18]. The problem of quantifying the code performance is often approached by examining the accuracy with which the stimulus can be ultimately decoded from the observed response, and the best decoding precision is evaluated using the Cramér-Rao lower bound on the mean squared error, see e.g. [19–24]. Since the probability distributions of the number of spikes are rarely known for advanced neuronal models, two approaches can be mostly encountered in the literature. Either processes based on a statistical description of the neuronal activity and enabling a straightforward way to express the count of spikes are used, e.g. Bernoulli or Poisson processes, [25–27], or the rate is inferred from interspike intervals [28].

In this work, the probability distribution of the number of spikes in a fixed time window for the stochastic perfect integrate-and-fire model is derived. We then use Fisher information to infer the rate coding accuracy, and investigate how it is influenced by the key parameters of the model. In particular, our first goal is to study the role of the observation time window. Throughout the paper, the time window is always the window used by the nervous system and not a window

determined by an external observer. Our second goal is to investigate how the amount of noise in the form of either presynaptic spontaneous activity or the random fluctuation of the membrane potential impacts the decoding accuracy, and if there is an optimal level of noise enhancing the signal, like for example in stochastic resonance and other similar phenomena, which were found in many settings (see, e.g., [29–36], and many more). As we show, counting spikes in a longer time window does not always improve the decoding accuracy. Moreover, the loss in decoding precision caused by a change in the length of the time window might be, at least partially, compensated by increasing the level of spontaneous activity of presynaptic neurons. Finally, we identify several stochastic-resonance-like phenomena related to the level of spontaneous activity and the level of fluctuation of the membrane potential.

II. METHODS

A. Fisher information

A common approach to analyze the decoding performance is to ask how well an optimal decoder can estimate the true value of the presented stimulus s based on a stochastic neuronal response R [26, 31, 37, 38]. To quantify the estimation accuracy, the mean squared error is used,

$$\mathrm{MSE}(\hat{s}) = \mathbb{E}\left[(\hat{s}-s)^2\right], \tag{1}$$

where $\hat{s}$ is the estimator of s from R. If the estimator $\hat{s}$ is unbiased, i.e. $\mathbb{E}(\hat{s}) = s$, the mean squared error is equal to the variance of the estimator, $\mathrm{MSE}(\hat{s}) = \mathrm{Var}(\hat{s})$. According to the Cramér-Rao theorem [39, 40], every unbiased and regular estimator $\hat{s}$ satisfies the inequality

$$\mathrm{Var}(\hat{s}) \geq \frac{1}{J(s)}, \tag{2}$$

where $J(s)$ denotes the Fisher information about s carried by R. For the sake of simplicity and analytical tractability, many theoretical studies on neural coding use the inverse of Fisher information as an approximation of the mean squared error, which implies that a higher Fisher information is assumed to reflect a higher decoding accuracy [29, 31].

From the rate coding perspective, the response R is the number of spikes after the stimulus onset. Hence, R is a discrete quantity, with probability mass function (pmf) $p_R(r;s)$ and the Fisher information is given by

$$J(s) = \sum_r \frac{1}{p_R(r;s)} \left[\frac{\partial}{\partial s} p_R(r;s)\right]^2. \tag{3}$$

Note that the stimulus level s plays the role of a parameter of the distribution of R.

B. Neuronal model

Throughout the study, we describe the neuronal activity by the stochastic perfect integrate-and-fire model introduced by Gerstein and Mandelbrot [41]. The idea of the model is to approximate the membrane potential dynamics by a Wiener process $X(t)$, which is given as the solution to a stochastic differential equation

$$dX(t) = \mu\, dt + \sigma\, dW(t), \qquad X(0) = 0,$$

where $W(t)$ is the standard (driftless) Wiener process, $\mu > 0$ is the drift and $\sigma > 0$ is the diffusion parameter. The membrane potential is thus driven by a deterministic force μ accompanied by white noise. Once $X(t)$ reaches a constant threshold $B > 0$, a spike is elicited, $X(t)$ is immediately reset to its starting value 0 and the accumulation of the membrane potential starts anew. The resulting spike train is then a renewal point process, where inter-spike intervals are independent and identically distributed as $T \sim IG(B/\mu, B^2/\sigma^2)$, an inverse Gaussian distribution with mean $\mathbb{E}(T) = B/\mu$ and variance $\mathrm{Var}(T) = B\sigma^2/\mu^3$.

Denote by X_0 the value of the membrane potential at a randomly chosen time. The distribution of X_0 can be derived from the known transition probability density of the Wiener process under the absorbing boundary B, which gives the probability density that the process starting from zero at time zero attains the value x at time t while the boundary B has not been crossed during that time [42, 43]

$$f_X^B(x,t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \left\{ \exp\left[-\frac{(x-\mu t)^2}{2\sigma^2 t} \right] - \exp\left[\frac{2\mu B}{\sigma^2} - \frac{(x-2B-\mu t)^2}{2\sigma^2 t} \right] \right\}. \tag{4}$$

The conditional density of X_0 given that the time elapsed since the last spike is t can be written as

$$f_{X_0|T>t}(x|t) = \frac{f_X^B(x,t)}{1 - F_T(t)}, \tag{5}$$

where $F_T(t)$ is the cumulative distribution function (cdf) of the inter-spike interval T. The time interval between the randomly chosen time point and the last preceding spike is the backward recurrence time S, which has the probability density function (pdf) $f_S(t) = [1 - F_T(t)]/\mathbb{E}(T)$ [44]. Multiplying the conditional density $f_{X_0|T>t}(x|t)$ by $f_S(t)$ and integrating over all possible t yields

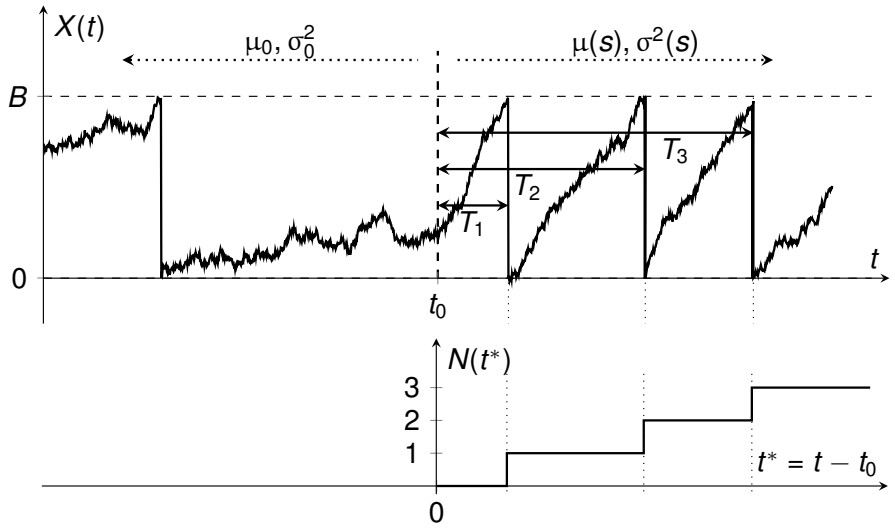

FIG. 1. Schematic illustration of the studied model. *Top:* A sample path of the membrane potential $X(t)$, described by a Wiener process. When $X(t)$ exceeds a constant threshold $B > 0$, an action potential is generated. Then $X(t)$ is reset to 0 and its evolution starts anew. At time t_0, a stimulus is applied, and the parameters of the process change from μ_0, σ_0^2 to $\mu(s)$, $\sigma^2(s)$. The times from the stimulus onset to the first, second, third,..., n-th evoked spike are denoted by T_1, T_2, T_3, ..., T_n. *Bottom:* The corresponding counting process of evoked spikes $N(t^*)$, which gives the number of spikes in a time window $[t_0, t_0 + t^*]$.

the unconditional distribution of X_0 as follows

$$f_{X_0}(x) = \int_0^\infty f_{X_0|T>t}(x|t) f_S(t)\, dt = \int_0^\infty \frac{f_X^B(x,t)}{\mathbb{E}(T)}\, dt$$
$$= \frac{1}{B}\left[\exp\left(\frac{\mu(x-|x|)}{\sigma^2}\right) - \exp\left(\frac{2\mu(x-B)}{\sigma^2}\right)\right]. \tag{6}$$

We consider a situation where a stimulus of level s is presented at time t_0. The spontaneous activity before t_0 results only from the spontaneous activity of presynaptic neurons, while the evoked activity after t_0 is affected by the stimulation. Before t_0, the parameters of the Wiener process are $\mu = \mu_0$ and $\sigma^2 = \sigma_0^2$. The presentation of the stimulus of intensity s at t_0 changes the parameters to $\mu = \mu(s)$ and $\sigma^2 = \sigma^2(s)$ (Fig. 1).

Throughout the paper we assume that the evoked drift $\mu(s)$ is a sum of the spontaneous drift μ_0 and the stimulus-driven increment $\Delta\mu(s)$. The specific functional form of $\mu(s)$ is derived from the Hill function [45], which is frequently employed in both theoretical and experimental studies

[46–50]. For a stimulus level s expressed on a logarithmic scale, it can be written as

$$\mu(s) = \mu_0 + \frac{A}{1+\mathrm{e}^{-b(s-s_0)}}, \qquad s \in (-\infty, \infty). \tag{7}$$

The parameter A is the maximum possible increment in the drift $\mu(s)$, b controls the steepness of the function and s_0 is the inflection point.

For the diffusion parameter $\sigma^2(s)$, we assume a linear dependence on the drift $\mu(s)$, that is

$$\sigma^2(s) = k\mu(s) + m, \tag{8}$$

where $k, m \geq 0$. The linear relationship $\sigma^2(s) = k\mu(s) + m$ can be derived for the membrane potential described by a randomized random walk with Poissonian input, where the amount of excitatory input is proportional to the stimulus intensity and the inhibitory input is constant [51, p. 138]. By letting either k or m be equal to zero, we obtain the following two special cases:

1. $\sigma^2(s) = k\mu(s)$.
   This corresponds to balanced excitatory and inhibitory input [52, 53].

2. $\sigma^2(s) = m$.
   Together with the requirement $\lim_{s\to-\infty} \sigma^2(s) = \sigma_0^2$, this gives a diffusion parameter independent of the stimulation, that is $\sigma^2(s) = \sigma_0^2$.

In the following we study the accuracy with which the stimulus intensity s can be decoded from $N(t^*)$, the number of spikes observed in a time window of duration t^*.

III. RESULTS

A. Fisher information

The reported Fisher information $J(s)$ about s based on the observation of $N(t^*)$ was calculated numerically using the formulas for the distribution of $N(t^*)$ given in Appendix. Throughout the paper, we consider two different scenarios corresponding to two possible beginnings of the observation time window:

1. *The observation time window starts with an evoked spike.*
   The initial value of the membrane potential is known and $X_0 = 0$. (Fig. 2 and 4*a*.)

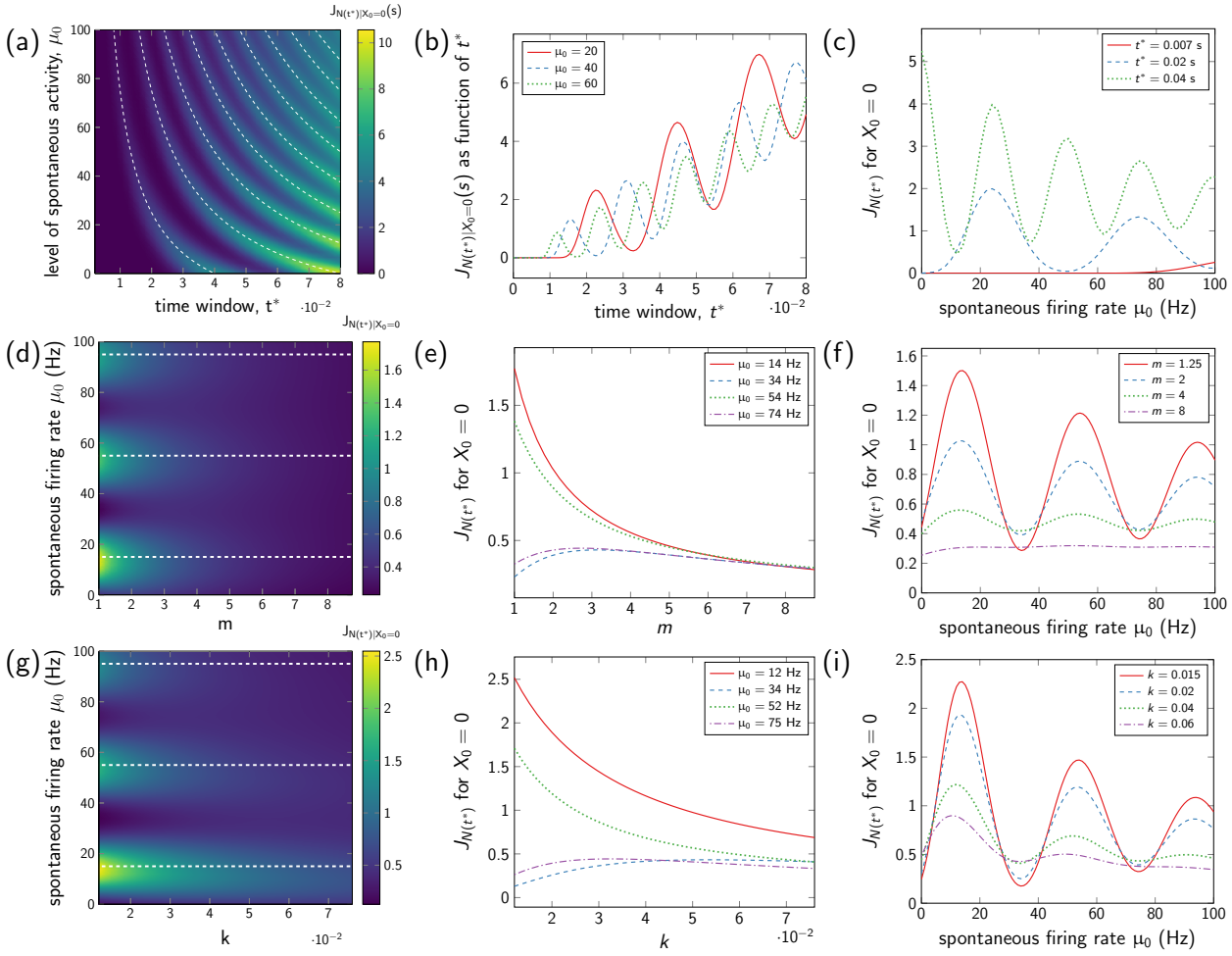

FIG. 2. Caption on the next page.

2. *The observation time window starts at t_0.*

   The initial value of the membrane potential X_0 is random and has pdf (6), with $\mu = \mu_0$ and $\sigma^2 = \sigma_0^2$. (Fig. 3 and 4*b*.)

We focus on the dependence of the Fisher information on the duration of the time window t^*, but we also take into account the level of spontaneous activity μ_0 and the parameters k and m governing the diffusion parameter $\sigma^2(s)$. Without loss of generality, the threshold B is set to $B = 1$ and, as a consequence, drift μ represents also the firing rate. Moreover, we set $A = 50$, $b = 1$ and $s_0 = 0$.

Fisher information with respect to the length of the time window t^* In general, Fisher information roughly increases as the time window gets longer (Fig. 2*a*, *b* when the observation time window starts with a spike, i.e. $X_0 = 0$, and 3*a*, *b* when the time window starts at t_0, i.e. X_0 is random). In both cases, however, the Fisher information is nonmonotone and has many local maxima and minima. The local extremes are more pronounced when $X_0 = 0$, but they are also clearly

FIG. 2. (Previous page.) Fisher information $J(0)$ about the stimulus intensity $s = 0$ based on the number of spikes $N(t^*)$ in a time window of length t^*. The observation time window starts with a spike and the initial value of the membrane potential at the beginning of the time window is thus $X_0 = 0$. In all cases, the diffusion coefficient is $\sigma^2(s) = k\mu(s) + m$. Panels (*a*), (*b*), (*c*): Fisher information as a function of the spontaneous firing rate μ_0 and the duration of the time window t^* when $k = 0.01$ and $m = 0.5$. The Fisher information is nonmonotone with respect to both μ_0 and t^*. Panels (*d*), (*e*), (*f*): Fisher information as a function of the spontaneous firing rate μ_0 and the absolute coefficient m of the diffusion parameter $\sigma^2(s)$ when $t^* = 0.025\,\mathrm{s}$ and $k = 0.01$. Panels (*g*), (*h*), (*i*): Fisher information as a function of the spontaneous firing rate μ_0 and the coefficient of proportionality k of the diffusion parameter $\sigma^2(s)$ when $t^* = 0.025\,\mathrm{s}$ and $m = 0.5$. For some values of μ_0 (dashed and dot-dashed lines in Panels *e, f*), the Fisher information is initially increasing with respect to k and m, and thus with respect to the fluctuation level of the membrane potential. The white dashed lines in the first column (Panels *a, d* and *g*) mark the points satisfying $t^* = nB/\mu(s)$, which approximately correspond to the local maxima of $J(s)$ with respect to t^*, where $B = 1$ is the membrane potential threshold and $n = 1, 2, \ldots$.

visible when X_0 is random. Therefore, observing the process for a longer time may not necessarily lead to a higher decoding accuracy.

Fisher information with respect to the level of spontaneous activity μ_0 When the response starts with a spike, the Fisher information has again a nonmonotone behavior with respect to μ_0, with many local maxima and minima (Fig. 2*a, c, d, f, g, i*) and a qualitatively similar behavior is also observed when X_0 is random, (Fig. 3*a, c, d, f, g, i*). Therefore, increasing the amount of presynaptic spontaneous activity may result in both improving or deteriorating the accuracy of decoding the stimulus, depending on the particular parameter conditions. For example, from Fig. 2*c* we observe that the Fisher information can be initially decreasing or increasing with respect to μ_0 depending on whether the observation time window is small ($t^* = 0.007\,\mathrm{s}$, $t^* = 0.02\,\mathrm{s}$) or large ($t^* = 0.04\,\mathrm{s}$).

If we ignore local extremes, the overall tendency is that the Fisher information decreases with respect to μ_0 due to the variability of the underlying process, which is linearly proportional to μ_0. When $k = 0$, however, the Fisher information has a mild overall increasing tendency (Fig. 4). This is because more spikes can be observed in the same time window due to the higher spontaneous activity, while the variability of the process remains constant (Fig. 4).

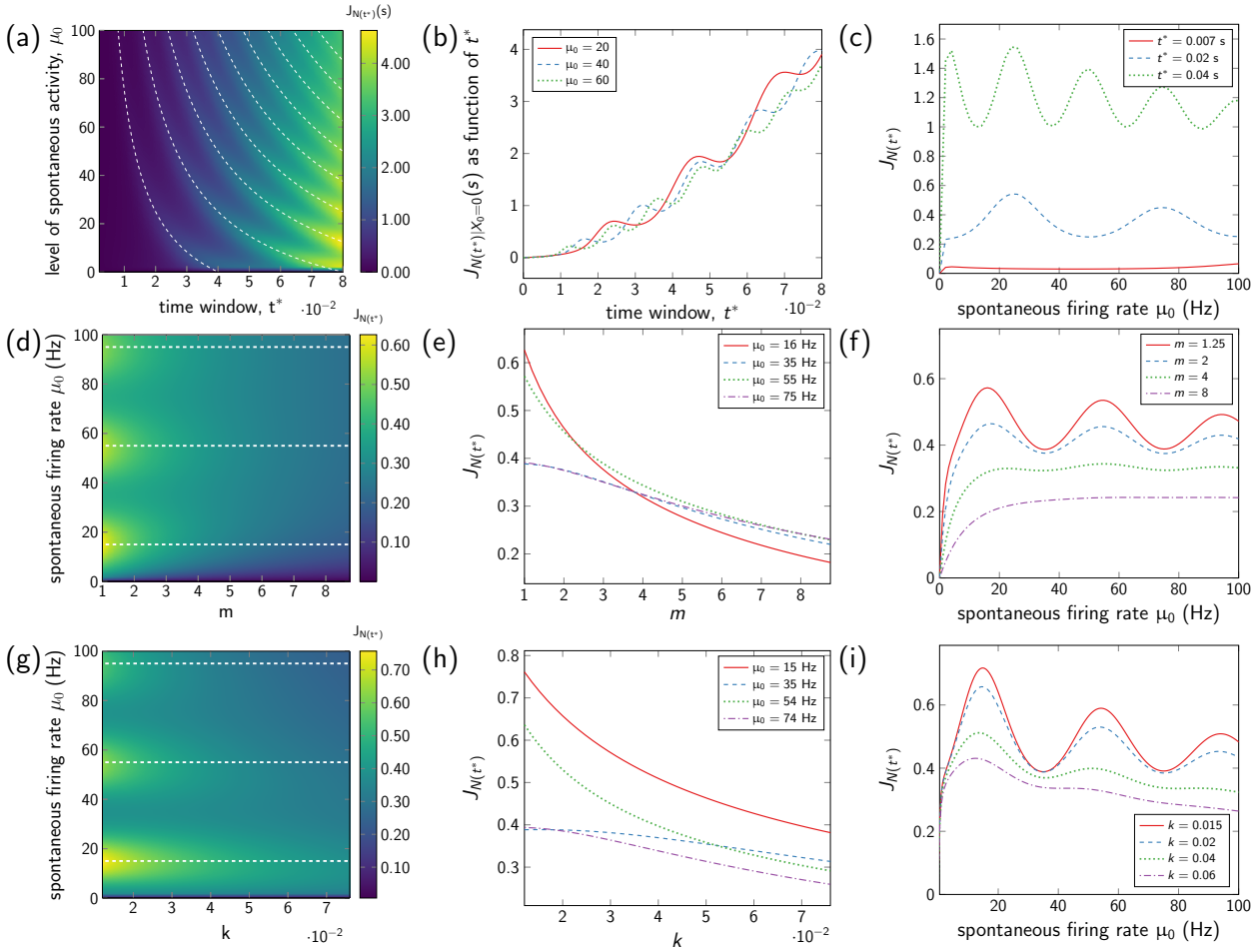

FIG. 3. Fisher information $J(0)$ about the stimulus intensity $s = 0$ when the time window starts at t_0 and the initial value of the membrane potential X_0 is thus random. In all cases, the coefficient diffusion is $\sigma^2(s) = k\mu(s) + m$. Panels (a), (b), (c): Fisher information as a function of the spontaneous firing rate μ_0 and the time window t^* when $k = 0.01$ and $m = 0.5$. The Fisher information is nonmonotone with respect to both μ_0 and t^*. Panels (d), (e), (f): Fisher information as a function of μ_0 and m when $t^* = 0.025\,\mathrm{s}$ and $k = 0.01$. Panels (g), (h), (i): Fisher information as a function of μ_0 and k when $t^* = 0.025\,\mathrm{s}$ and $m = 0.5$. For any choice of μ_0, the Fisher information is decreasing in k and m. The white dashed lines in the first column (Panels *a, d* and *g*) mark the points satisfying $t^* = nB/\mu(s)$, which approximately correspond to the local maxima of the Fisher information with respect to t^*, where $B = 1$ is the membrane potential threshold and $n \in \mathbb{N}$.

It is interesting to look at the mutual effect of both the spontaneous activity and the time window on the Fisher information, see Fig. 2*a* and Fig. 3*a* when $X_0 = 0$ and X_0 is random, respectively. The white dashed lines displayed in the figures mark all points where the time window is $t^* = nB/\mu(s)$, $n \in \mathbb{N}$, which approximately correspond to the local maxima of the Fisher information with respect

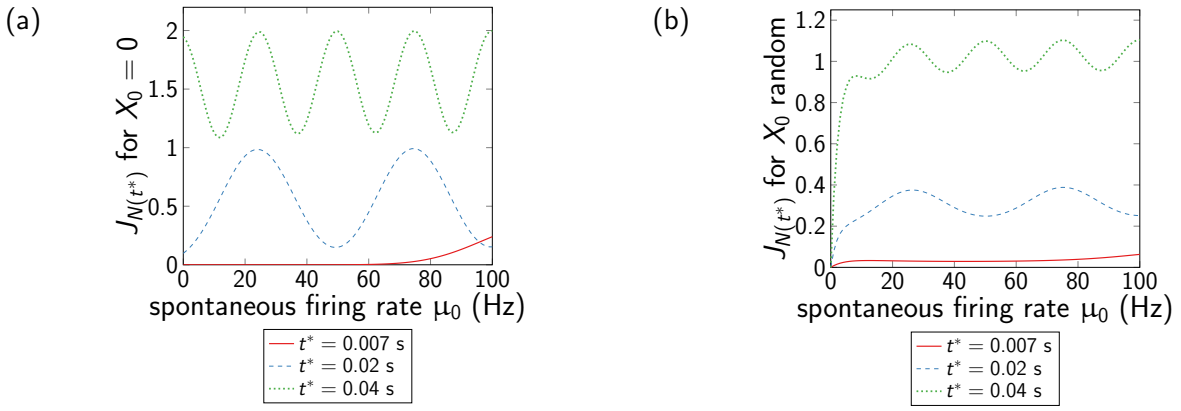

FIG. 4. Fisher information $J(0)$ about the stimulus intensity $s = 0$ as a function of the spontaneous firing rate μ_0 when the diffusion coefficient is independent of stimulation, namely $\sigma^2(s) = \sigma_0^2 = 2$. (*a*) Fisher information when the observation time window starts with a spike. (*b*) Fisher information when the observation window starts at t_0. Three lengths of the observation window are used: $t^* = 0.007\,\mathrm{s}$ (red, solid), $t^* = 0.02\,\mathrm{s}$ (blue, dashed), $t^* = 0.04\,\mathrm{s}$ (green, dotted). If the local extremes are ignored, the Fisher information has a mild overall increasing tendency in all the cases, suggesting that the decoding accuracy improves with increasing spontaneous firing rate μ_0.

to t^* for fixed μ_0 derived when $\sigma_0^2 = \sigma^2(s) = 0$, as discussed in more detail in Section III B. If the parameters of the system are chosen to be near a local maximum of $J(s)$ with respect to t^* and then the observation time window is taken shorter, the loss in the decoding accuracy can be to some extent compensated by increasing the level of spontaneous activity μ_0. Similarly, if the level of spontaneous activity is increased, by shortening the observation window we may get almost the original decoding accuracy.

Fisher information with respect to the diffusion coefficient $\sigma^2(s)$ If X_0 is random, increasing the diffusion parameter $\sigma^2(s)$ through increasing k or m always yields a lower Fisher information (Fig. 3*d-i*). On the other hand, when $X_0 = 0$, in some cases the Fisher information is initially increasing with respect to k and m (Fig. 2*d-i*), meaning that a positive level of noise enhances the signal. There are also cases, though, where the Fisher information is strictly decreasing with respect to k or m. Whether $J(s)$ is increasing or decreasing in a given situation depends on the level of spontaneous activity. If the spontaneous drift μ_0 is approximately around

$$\mu_0 \approx \frac{nB}{t^*} - \frac{A}{1+\exp(-b(s-s_0))}, \qquad n \in \mathbb{N}, \tag{9}$$

then increasing $\sigma^2(s)$ has the effect of deteriorating the decoding accuracy. On the other hand, if the spontaneous drift μ_0 is around

$$\mu_0 \approx \frac{2n(n+1)B}{(2n+1)t^*} - \frac{A}{1+\exp(-b(s-s_0))}, \qquad n \in \mathbb{N} \tag{10}$$

and $\sigma^2(s)$ is small, then increasing $\sigma^2(s)$ through k or m may improve the decoding accuracy. Finally, increasing k or m causes that the local extremes of the Fisher information with respect to t^* and μ_0 become smaller and for a sufficiently large $\sigma^2(s)$ they are almost negligible.

B. Comparison with the limit case $\sigma^2 = 0$

The counter-intuitive nonmonotone behavior of the Fisher information can be explained by looking at the limit case $\sigma_0^2 = \sigma^2(s) = 0$, when the process $X(t)$ is deterministic. The only randomness in the system is then the initial value of the membrane potential X_0 when the observation time window does not start with a spike.

Denote by T_n the time from t_0 to the n-th following spike. If $X_0 = x_0$ is known, $T_n = t_n$ is deterministic and equal to

$$t_n = (nB - x_0)/\mu(s). \tag{11}$$

If X_0 is random, the cdf of T_n is given by

$$F_{T_n}(t) = \mathbb{P}\left(\frac{nB - X_0}{\mu(s)} \le t\right) = \mathbb{P}\big(X_0 \ge nB - \mu(s)t\big) = 1 - F_{X_0}(nB - \mu(s)t), \tag{12}$$

where F_{X_0} is the cdf of X_0. Consequently, we can write

$$p_{N(t^*)}(n) = \mathbb{P}(T_n \le t^* < T_{n+1}) = \begin{cases} F_{X_0}(B - \mu(s)t^*) & n = 0 \\ F_{X_0}\big((n+1)B - \mu(s)t^*\big) - F_{X_0}\big(nB - \mu(s)t^*\big) & n \ge 1. \end{cases} \tag{13}$$

If $\sigma_0^2 = 0$, X_0 is uniformly distributed over the interval $[0, B]$, so we obtain

$$p_{N(t^*)}(n) = \begin{cases} 0 & t^* \in \left(0, \frac{(n-1)B}{\mu(s)}\right) \\ 1 - n + \frac{\mu(s)t^*}{B} & t^* \in \left[\frac{(n-1)B}{\mu(s)}, \frac{nB}{\mu(s)}\right) \\ n + 1 - \frac{\mu(s)t^*}{B} & t^* \in \left[\frac{nB}{\mu(s)}, \frac{(n+1)B}{\mu(s)}\right) \\ 0 & t^* \in \left[\frac{(n+1)B}{\mu(s)}, \infty\right), \end{cases} \tag{14}$$

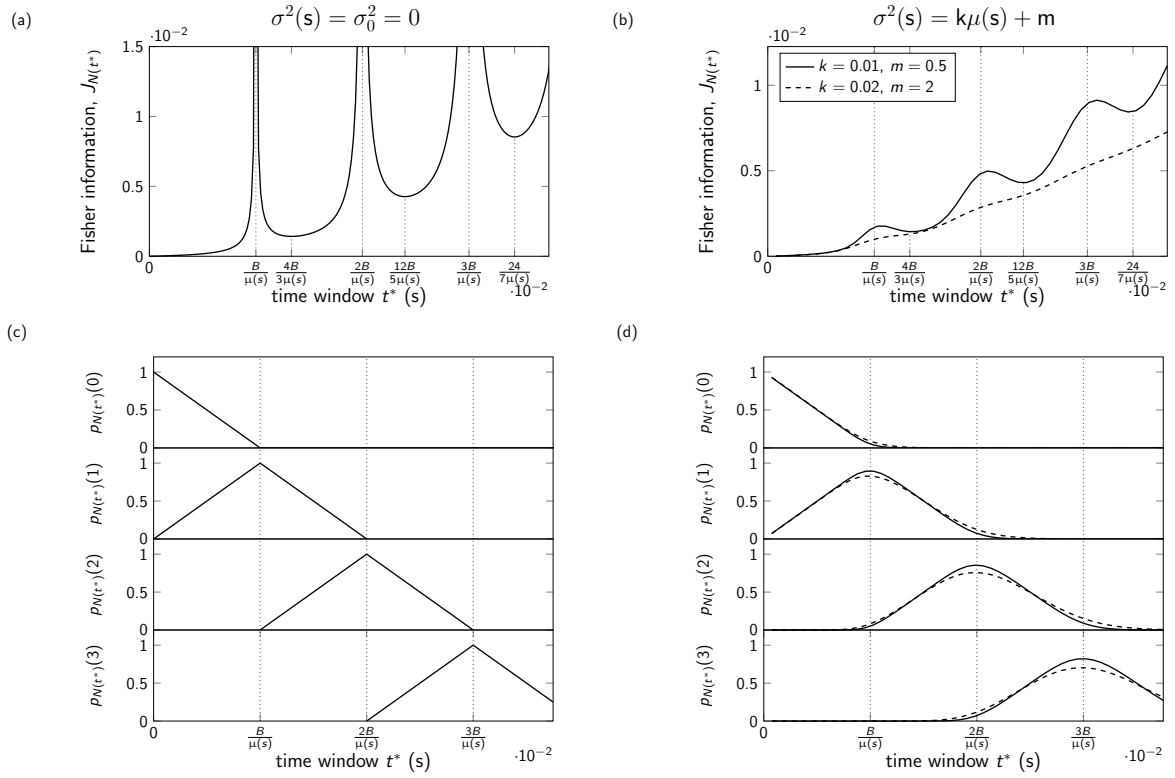

FIG. 5. (*a*) Fisher information $J(0)$ about the stimulus intensity $s = 0$ as a function of the length of the time window t^* for the limit case of zero diffusion coefficient, $\sigma_0^2 = \sigma^2(s) = 0$. The Fisher information goes to infinity for all t^* satisfying $t^* = nB/\mu(s)$, $n \in \{1,2,\dots\}$. (*b*) Fisher information $J(0)$ when $\sigma_0^2 > 0$ and $\sigma^2(s) > 0$. The local maxima are located approximately at $t^* = nB/\mu(s)$. (*c*) Probabilities of observing n spikes $p_{N(t^*)}(n) = \mathbb{P}(N(t^*) = n)$ as functions of t^* for the limit case $\sigma_0^2 = \sigma^2(s) = 0$. At $t^* = nB/\mu(s)$, the probability of observing n spikes is equal to 1 and all the other outcomes have zero probability. (*d*) The probabilities $p_{N(t^*)}(n)$ as functions of t^* when $\sigma_0^2 > 0$ and $\sigma^2(s) > 0$. At $t^* = nB/\mu(s)$, the probability of observing n spikes is nearly one and decreases when the diffusion coefficient $\sigma^2(s)$ is increased. All figures were obtained for the spontaneous firing rate $\mu_0 = 50$ Hz.

for $n \in \{1,2,\dots\}$, and for $n = 0$ we have

$$p_{N(t^*)}(0) = \begin{cases} 1 - \frac{\mu(s)t^*}{B} & t^* \in \left(0, \frac{B}{\mu(s)}\right) \\ 0 & t^* \in \left[\frac{B}{\mu(s)}, \infty\right). \end{cases} \tag{15}$$

The Fisher information about s based on observing $N(t^*)$ can be calculated as

$$J(s) = \frac{[\mu'(s)t^*]^2}{[(n+1)B - \mu(s)t^*][-nB + \mu(s)t^*]}, \qquad \text{if } t^* \in \left(\frac{nB}{\mu(s)}, \frac{(n+1)B}{\mu(s)}\right) \tag{16}$$

and is illustrated in Fig. 5*a*. It is a piecewise function that is defined on every interval $(nB/\mu(s),(n+1)B/\mu(s))$, $n=0,1,2,\ldots$ and for $t^*=nB/\mu(s)$, $n\in\{1,2,\ldots\}$ goes to infinity. The minimum Fisher information on the interval $t^*\in(nB/\mu(s),(n+1)B/\mu(s))$ is achieved at

$$t^*_{\min}=\frac{2n(n+1)B}{(2n+1)\mu(s)},\qquad n\in\mathbb{N}. \tag{17}$$

The discontinuities are caused by the shape of the probabilities $p_{N(t^*)}(n)=\mathbb{P}(N(t^*)=n)$, illustrated in Fig. 5*c*. The probabilities $p_{N(t^*)}(n)$ are piecewise linear and the derivatives with respect to s are not defined at $t^*=nB/\mu(s)$, $n\in\mathbb{N}$. At these points also $p_{N(t^*)}(n-1)$ and $p_{N(t^*)}(n+1)$ tend to zero and the corresponding terms in the formula for the Fisher information tend to infinity. At $t^*=nB/\mu(s)$, the probability of observing n spikes is equal to one.

Now we compare the Fisher information $J(s)$ for the limit case $\sigma_0^2=\sigma^2(s)=0$ (Fig. 5*a*) with the Fisher information $J(s)$ for the previously studied cases $\sigma_0^2>0$ and $\sigma^2(s)>0$ (Fig. 5*b*). It can be seen that the local maxima of the Fisher information for $\sigma_0^2>0$ and $\sigma^2(s)>0$ are located near $t^*=nB/\mu(s)$, $n\in\mathbb{N}$. This can also be confirmed in Fig. 2*a, d, g* and Fig. 3*a, d, g*, where the white dashed lines, corresponding to $t^*=nB/\mu(s)$ are very close to the local maxima of the Fisher information with respect to t^*. In Fig. 5*b* we can also notice that if the value of $\sigma^2(s)$ is increased by increasing k or m, the local maxima have the tendency to vanish.

We can also see the correspondence between the shape of the Fisher information $J(s)$ and the probabilities $p_{N(t^*)}(n)$. As $\sigma^2(s)$ increases, the probabilities $p_{N(t^*)}(n)$ become smoother with not so sharp transitions between the increasing and the decreasing part. Also the maximum of the probability $p_{N(t^*)}(n)$ is less than 1 (Fig. 5*d*), which results in smaller fluctuations of the Fisher information.

We can conclude that the nonmonotone behavior of the Fisher information is caused by the discrete character of $N(t^*)$. If the randomness of $X(t)$ in the form of $\sigma^2(s)$ is not too high, there are only a few possible outcomes of $N(t^*)$. Hence the Fisher information is a sum of only a small number of nonzero terms in the form $(\partial p_{N(t^*)}(n)/\partial s)^2/p_{N(t^*)(n)}$. Therefore, a big change in any of them is strongly reflected in the Fisher information. Note that an individual term has a big contribution if the respective probability is close to zero, which typically happens around $t^*=nB/\mu(s)$. The local maxima and minima of the Fisher information with respect to t^* appear periodically with the period approximately equal to $B/\mu(s)$, which is the mean inter-spike interval. Hence, this phenomenon is not restricted only to short time windows.

IV. DISCUSSION

This analysis reveals that even in neuronal models as simple as the perfect integrate-and-fire model the effect of spontaneous activity, noise and the length of the observation window might be highly nontrivial. In spite of the relative simplicity, the question of neuronal coding is often studied on this or a similar level of complexity [54–58], because the perfect integrate-and-fire model is the most complex setup allowing analytical calculations. Therefore, it is possible to identify the reasons for the nonmonotone behavior of the Fisher information. Since the perfect integrate-and-fire model is based on Brownian motion, it can be also applied on a wide variety of physical problems beyond the field of neuroscience [59, 60], implying that the described phenomena might appear also there.

Although analogous calculations for more detailed biophysical models, such as models including leakage or setups considering the activity of a group of neurons, which are the directions in which the analysis is to be extended in the future, are technically prohibitive, this fact cannot be used to imply that the described effects do not occur therein. On the contrary, the most important result presented here, i.e. that the Fisher information is nonmonotone with respect to the observation window t^*, is caused by the discrete nature of the counting distribution. Thus it is not specific for the chosen model and should be largely independent of the amount of biophysical detail taken into account. Our expectation concerning the role of μ_0, k and m in leaky integrate-and-fire models is that they could have even stronger effect, since they may help to move the process from subthreshold to suprathreshold regime (in case of μ_0) or increase the chance that $X(t)$ crosses the threshold in subthreshold regime (in case of k and m), which is the classical phenomenon of stochastic resonance [30].

We assumed one particular form of $\mu(s)$ given by the logistic function. However, the observed phenomena are more general. When analyzing the Fisher information with respect to t^* and k and m, we fix μ_0 and s, so the value of $\mu(s)$ is constant and the functional form of $\mu(s)$ plays no role. When we study the behavior of the Fisher information with respect to μ_0, the assumed form of $\mu(s)$ might have some influence. Nevertheless, our results with respect to μ_0 can also be obtained for other types of $\mu(s)$, for example for the linear relationship $\mu(s) = \mu_0 + s$, $s \geq 0$.

V. CONCLUSIONS

We studied the accuracy of the rate code with respect to several key parameters of the neuronal system and showed that the relationships are nontrivial. Altogether, our results may be summarized in the following four points:

1. Using a longer time window may not necessarily improve the decoding accuracy.

2. Although the presynaptic spontaneous activity might seem a disturbing element, since it bears no information about the stimulus, its presence at a certain level might improve the decoding accuracy.

3. The loss in the decoding accuracy caused by the change of the time window can be partially compensated if it is accompanied by an appropriate change of the level of spontaneous activity, and vice versa. For example, a shorter time window may sometimes give almost the same accuracy if the level of spontaneous activity is increased by an optimal amount.

4. If the time window begins with a spike, the decoding accuracy might be improved by increasing the fluctuation of the membrane potential.

ACKNOWLEDGEMENTS

This work was supported by the Joint Research Project between Austria and Czech Republic, 7AMB15AT010, and by the Czech Science Foundation project 15-08066S.

Appendix: Distribution of $N(t^*)$

Here we show how the distribution of $N(t^*)$ can be derived. The first step to obtain the distribution of $N(t^*)$ is to find the distribution of T_n, the time from t_0 to the n-th following spike.

a. Distribution of T_n Since the Wiener process is a time and space homogeneous process, the distribution of T_n, the n-th passage time of $X(t)$ to a constant threshold B, is the same as the first passage time of the same process $X(t)$ to the threshold nB. Thus, $T_n|X_0$, the time to the n-th spike given a fixed starting position $X_0 = x_0$, follows an inverse Gaussian distribution

$IG((nB-x_0)/\mu(s),(nB-x_0)^2/\sigma^2(s))$ with cdf

$$F_{T_n|X_0}(t|x_0)=\Phi\left(-\frac{nB-x_0-\mu(s)t}{\sigma(s)\sqrt{t}}\right)+\exp\left(\frac{2(nB-x_0)\mu(s)}{\sigma^2(s)}\right)\Phi\left(-\frac{nB-x_0+\mu(s)t}{\sigma(s)\sqrt{t}}\right). \quad (A.1)$$

Multiplying the pdf of $T_n|X_0$ by $f_{X_0}(x)$, given by (6) with parameters μ_0 and σ_0^2, and integrating over all possible x_0 values, we obtain the unconditional pdf of T_n

$$\begin{aligned}
f_{T_n}(t) &= \int_{-\infty}^{B} f_{T_n|X_0}(t|x)f_{X_0}(x)\,dx \\
&= \frac{\mu(s)\sigma_0^2-2\mu_0\sigma^2(s)}{B\sigma_0^2}\exp\left(\frac{2\mu_0(\mu_0\sigma^2(s)t-\mu(s)\sigma_0^2t+(n-1)B\sigma_0^2)}{\sigma_0^4}\right)\times \\
&\quad\times\left[\exp\left(\frac{2\mu_0 B}{\sigma_0^2}\right)\Phi\left(-\frac{nB\sigma_0^2+2\mu_0\sigma^2(s)t-\mu(s)\sigma_0^2t}{\sigma_0^2\sigma(s)\sqrt{t}}\right)-\right. \\
&\quad\left.-\Phi\left(-\frac{(n-1)B\sigma_0^2+2\mu_0\sigma^2(s)t-\mu(s)\sigma_0^2t}{\sigma_0^2\sigma(s)\sqrt{t}}\right)\right]+ \\
&\quad+\frac{\mu(s)}{B}\left[\Phi\left(\frac{nB-\mu(s)t}{\sigma(s)\sqrt{t}}\right)-\Phi\left(\frac{(n-1)B-\mu(s)t}{\sigma(s)\sqrt{t}}\right)\right]. \quad (A.2)
\end{aligned}$$

When $n=1$, the distribution of T_1 is identical to the distribution of the first-spike latency reported in [61].

b. Distribution of $N(t^*)$ Consider first the case when the membrane potential at time t_0 is fixed and known, i.e. $X_0=x_0$. The conditional probability of observing n spikes in the time window of length t^* given that $X_0=x_0$ can be written as

$$\begin{aligned}
p_{N(t^*)|X_0}(n|x_0) &= \mathbb{P}(N(t^*)=n|X_0=x_0) \\
&= \mathbb{P}(T_n\le t^*|X_0=x_0)-\mathbb{P}(T_{n+1}\le t^*|X_0=x_0). \quad (A.3)
\end{aligned}$$

Plugging (A.1) into (A.3), we obtain

$$\begin{aligned}
p_{N(t^*)|X_0}(n|x_0) &= \Phi\left(-\frac{nB-x_0-\mu(s)t^*}{\sigma(s)\sqrt{t^*}}\right)-\Phi\left(-\frac{(n+1)B-x_0-\mu(s)t}{\sigma(s)\sqrt{t^*}}\right)+ \\
&\quad+\exp\left(\frac{2\mu(s)[nB-x_0]}{\sigma^2(s)}\right)\left[\Phi\left(-\frac{nB-x_0+\mu(s)t^*}{\sigma(s)\sqrt{t^*}}\right)-\right. \\
&\quad\left.-\exp\left(\frac{2\mu(s)B}{\sigma^2(s)}\right)\Phi\left(-\frac{(n+1)B-x_0+\mu(s)t^*}{\sigma(s)\sqrt{t^*}}\right)\right] \quad (A.4)
\end{aligned}$$

for $n=1,2,\ldots$. If $n=0$, we have

$$p_{N(t^*)|X_0}(0|x_0)=1-\Phi\left(-\frac{B-x_0-\mu(s)t^*}{\sigma(s)\sqrt{t^*}}\right)-\exp\left(\frac{2(B-x_0)\mu(s)}{\sigma^2(s)}\right)\Phi\left(-\frac{B-x_0+\mu(s)t^*}{\sigma(s)\sqrt{t^*}}\right). \quad (A.5)$$

If the value of X_0 is not known, the straightforward way to get the unconditional density of $N(t^*)$ is to calculate the joint density of $N(t^*)$ and X_0 and integrate over all possible values of X_0, that is

$$p_{N(t^*)}(n) = \int_{-\infty}^{B} \mathbb{P}(N(t^*) = n | X_0 = x) f_{X_0}(x)\, dx \tag{A.6}$$

using (A.4) and (6). Unfortunately, the closed form of (A.6) cannot be obtained and numerical integration must be done. To avoid possible numerical issues, we suggest to compute $p_{N(t^*)}(n)$ as

$$p_{N(t^*)}(n) = \begin{cases} \int_0^{t^*} f_{T_n}(\tau)[1 - F_{T_1}(t-\tau)]\, d\tau & n \geq 1 \\ 1 - F_{T_1}(t^*) & n = 0, \end{cases} \tag{A.7}$$

which has the numerical advantage of performing the numerical integration only over a finite interval.

---

[1] E. D. Adrian, *The basis of sensation* (WW Norton & Co, 1928).

[2] W. Gerstner and W. M. Kistler, *Spiking Neuron Models* (Cambridge University Press, 2002).

[3] F. Theunissen and J. P. Miller, J. Comput. Neurosci. **2**, 149 (1995).

[4] R. Kobayashi and S. Shinomoto, Phys. Rev. E **75**, 011925 (2007).

[5] Y. Mochizuki and S. Shinomoto, Phys. Rev. E **89**, 022705 (2014).

[6] W. Bialek, F. Rieke, R. R. De Ruyter Van Steveninck, and D. Warland, Science **252**, 1854 (1991).

[7] M. N. Shadlen and W. T. Newsome, Curr. Opin. Neurobiol. **4**, 569 (1994).

[8] M. Oram, M. Wiener, R. Lestienne, and B. Richmond, J. Neurophysiol. **81**, 3021 (1999).

[9] A. L. Jacobs, G. Fridman, R. M. Douglas, N. M. Alam, P. E. Latham, G. T. Prusky, and S. Nirenberg, Proceedings of the National Academy of Sciences **106**, 5936 (2009).

[10] M. Ainsworth, S. Lee, M. O. Cunningham, R. D. Traub, N. J. Kopell, and M. A. Whittington, Neuron **75**, 572 (2012).

[11] S. Thorpe, D. Fize, C. Marlot, *et al.*, Nature **381**, 520 (1996).

[12] J. Feng, Journal of Physics A: Mathematical and General **34**, 7475 (2001).

[13] S. Wu, S.-i. Amari, and H. Nakahara, Neural Comput. **14**, 999 (2002).

[14] S. D. Wilke and C. W. Eurich, Neural Comput. **14**, 155 (2002).

[15] N. Gordon, T. M. Shackleton, A. R. Palmer, and I. Nelken, J. Neurosci. Methods **169**, 391 (2008).

[16] F. Montani, E. Phoka, M. Portesi, and S. R. Schultz, Physica A **392**, 3066 (2013).

[17] M. Deger, T. Schwalger, R. Naud, and W. Gerstner, Phys. Rev. E **90**, 062704 (2014).

[18] S. Koyama and L. Kostal, Math Biosci Eng **11**, 63 (2014).

[19] P. Dayan and L. F. Abbott, *Theoretical neuroscience* (MIT Press, 2001).

[20] N. Brunel and J.-P. Nadal, Neural Comput **10**, 1731 (1998).

[21] I. Dean, N. S. Harper, and D. McAlpine, Nature Neurosci. **8**, 1684 (2005).

[22] L. Kostal, P. Lansky, and S. Pilarski, J. Neural Eng. **12**, 036012 (2015).

[23] S. Pilarski and O. Pokora, BioSystems **136**, 11 (2015).

[24] R. E. Spinney, J. T. Lizier, and M. Prokopenko, Phys. Rev. E **94**, 022135 (2016).

[25] R. L. Winslow and M. B. Sachs, Hearing Res. **35**, 165 (1988).

[26] M. Bethge, D. Rotermund, and K. Pawelzik, Neural Comput **14**, 2317 (2002).

[27] D. H. Johnson and W. Ray, J Comput Neurosci **16**, 129 (2004).

[28] P. Lansky and P. E. Greenwood, BioSystems **89**, 10 (2007).

[29] M. Stemmler, Network **7**, 687 (1996).

[30] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, Reviews of Modern Physics **70**, 223 (1998).

[31] P. E. Greenwood, L. M. Ward, and W. Wefelmeyer, Phys. Rev. E **60**, 4687 (1999).

[32] F. Chapeau-Blondeau, S. Blanchard, and D. Rousseau, Phys. Rev. E **74**, 031102 (2006).

[33] L. A. Gatys, A. S. Ecker, T. Tchumatchenko, and M. Bethge, Phys. Rev. E **91**, 062707 (2015).

[34] I. Bashkirtseva, A. B. Neiman, and L. Ryashko, Phys. Rev. E **91**, 052920 (2015).

[35] J. Zylberberg and E. Shea-Brown, Phys. Rev. E **92**, 062707 (2015).

[36] M. Levakova, M. Tamborrino, L. Kostal, and P. Lansky, Neural Comput. **28**, to appear (2016).

[37] S. Amari and H. Nakahara, Neural Comput. **17**, 839 (2005).

[38] P. E. Greenwood and P. Lansky, Biol. Cybern. **92**, 199 (2005).

[39] E. J. G. Pitman, *Some basic theory for statistical inference* (Chapman and Hall, 1979).

[40] C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed. (John Wiley & Sons, New York, 2002).

[41] G. Gerstein and B. Mandelbrot, Biophys. J. **4**, 41 (1964).

[42] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes* (Chapman and Hall, 1965).

[43] M. T. Giraudo, P. E. Greenwood, and L. Sacerdote, Neural Comput. **23**, 1743 (2011).

[44] D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events* (Methuen, 1966).

[45] S. A. Frank, Biology Direct **8**, 31 (2013).

[46] M. Chastrette, T. Thomas-Danguin, and E. Rallet, Chem Senses **23**, 181 (1998).

[47] L. Nizami, Hearing Res. **167**, 13 (2002).

[48] J.-P. Rospars, P. Lansky, A. Duchamp, and P. Duchamp-Viret, Eur. J. Neurosci. **18**, 1135 (2003).

[49] S. Durant, C. W. G. Clifford, N. A. Crowder, N. S. C. Price, and M. R. Ibbotson, J Opt. Soc. Am. A **24**, 1529 (2007).

[50] A. Grémiaux, T. Nowotny, D. Martinez, P. Lucas, and J.-P. Rospars, Brain Res. **1434**, 123 (2012).

[51] H. C. Tuckwell, *Introduction to Theoretical Neurobiology, Vol.2: Nonlinear and Stochastic Theories* (Cambridge Univ. Press, Cambridge, 1988).

[52] K. Miura, Y. Tsubo, M. Okada, and T. Fukai, J. Neurosci. **27**, 13802 (2007).

[53] B. Sengupta, S. B. Laughlin, and J. E. Niven, PLoS Comput. Biol. **9**, e1003263 (2013).

[54] B. Lindner, Phys Rev E **69**, 022901 (2004).

[55] D. Andrieux and T. Monnai, Phys Rev E **80**, 021933 (2009).

[56] R. D. Vilela and B. Lindner, Phys Rev E **80**, 031909 (2009).

[57] F. Müller-Hansen, F. Droste, and B. Lindner, Phys Rev E **91**, 022718 (2015).

[58] R. Mankin and A. Rekker, Phys Rev E **94**, 062103 (2016).

[59] R. Mankin, K. Laas, and N. Lumi, Phys Rev E **88**, 042142 (2013).

[60] R. Mankin, K. Laas, N. Lumi, and A. Rekker, Phys Rev E **90**, 042127 (2014).

[61] M. Tamborrino, S. Ditlevsen, and P. Lansky, Lifetime Data Anal. **21**, 331 (2015).