

Optimal noise-aided signal transmission through populations of neurons

Thomas Hoch, Gregor Wenning, and Klaus Obermayer

Department of Electrical Engineering and Computer Science, Technical University of Berlin, Franklinstraße 28/29, 10587 Berlin, Germany

(Received 8 April 2003; published 29 July 2003)

Metabolic considerations and neurophysiological measurements indicate that biological neural systems prefer information transmission via many parallel low intensity channels, compared to few high intensity ones [S. B. Laughlin *et al.*, *Nature Neurosci.* **1**, 36 (1998)]. Furthermore, cortical neurons are exposed to a considerable amount of synaptic background activity, which increases the neurons' conductance and leads to a fluctuating membrane potential that, on average, is close to the threshold [A. Destexhe and D. Paré, *J. Neurophysiol.* **81**, 1531 (1999)]. Recent studies have shown that noise can improve the transmission of subthreshold signals in populations of neurons, e.g., if their response is pooled. In general, the optimal noise level depends on the stimulus distribution and on the number of neurons in the population. In this contribution we show that for a large enough number of neurons the latter dependency becomes weak, such that the optimal noise level becomes almost independent of the number of neurons in the population. First we investigate a binary threshold model of neurons. We derive an analytic expression for the optimal noise level at each single neuron, which—for a large enough population size—depends only on quantities that are locally available to a single neuron. Using numerical simulations, we then verify the weak dependence of the optimal noise level on population size in a more realistic framework using leaky integrate-and-fire as well as Hodgkin-Huxley-type model neurons. Next we construct a cost function, where quality of information transmission is traded against its metabolic costs. Again we find that—for subthreshold signals—there is an optimal noise level which maximizes this cost. This noise level, however, is almost independent of the number of neurons, even for small population sizes, as numerical simulations using the Hodgkin-Huxley model show. Since the dependence of the optimal noise level on population size is weak for large enough populations, local neural adaptation is sufficient to adjust the level of noise to its optimal value.

DOI: 10.1103/PhysRevE.68.011911

PACS number(s): 87.18.Sn, 87.10.+e, 87.19.La

I. INTRODUCTION

A basic feature of spike trains observed in many experiments *in vivo* is the high degree of variability, which either may be due to intrinsic noise sources or may be the signature of a probabilistic code. Recent studies have suggested that this irregular firing arises from the background activity, which is permanently present in the cortex [1]. This fluctuating activity can significantly influence the information transmission properties of cortical neurons by making transmission fast [2] or by allowing the transmission of subthreshold signals, for example, in a stochastic resonance setting [3]. Here we investigate the role of noise on the dynamics of subthreshold input signals. When a noisy current is injected into the neuron, stochastic resonance (see Ref. [4] for a review) occurs and the transmission of subthreshold input signals is enhanced [5–7]. Information transmission—measured, for example, by the mutual information between input and output—becomes optimal for a certain noise level, which depends on the properties of the distribution of the input signal. Recently, it has been shown for a variety of neuron models that the optimal noise level is linked to an optimal output rate [8], and it has been suggested that neurons may adjust their individual noise level based on quantities that are accessible to a single neuron and its elements.

In the central nervous system of higher animals, however, single neurons rarely matter, and information is likely to be coded using populations of cells [9]. For the cerebral cortex of higher animals, population size has been estimated to be

between 100 and 200 neurons [10], but this number may differ between brain areas and species. Hence the question arises, whether noise may aid the transmission of weak signals also through populations of neurons, and what quantities optimal signal transmission depends on. Recently, Collins *et al.* [11] and Stocks and Mannella [12] have examined the information transmission properties of a summing network of FitzHugh-Nagumo model neurons. They showed that stochastic resonance occurs and that there exists an optimal noise level for signal transmission. If the dependency on the size would be strong, local adaptation rules—as suggested in Ref. [8]—might not suffice, because information about population size would have to be made available to the single neuron.

In the paper of Collins *et al.* it was shown that stochastic resonance curves become broad if the number of neurons in the population becomes large. This is beneficial for a neural system, because it alleviates the above mentioned problem of how a single neuron can locally adapt its optimal noise level to different input distributions and to the correct population size. However, the study of Collins *et al.* is based on a cross-correlation measure, which is bounded from above and which may lead to a somewhat “squeezed” stochastic resonance curve when compared to other measures of information transmission. Both studies are also restricted to the FitzHugh-Nagumo model, and the question remains, how dependent this effect is on the particular mathematical description. This motivates a more detailed investigation using different measures of information transmission and a larger

variety of neuron models, including models which include the role of varying membrane conductances.

Another issue to be considered is that maximal information transmission may not be the only goal neurons try to achieve. Neural activity is costly in metabolic terms, and energy consumption and dissipation becomes a concern, for example, for the densely packed central nervous system of higher animals. Several researchers have suggested that the overall energy consumption constrains information transmission, and it has been argued that neurons try to achieve a balance between information transmission and energy consumption, leading to energy efficient codes [13–16]. Energy efficient codes favor low spike rates and subthreshold input distributions [14], which raises the hypothesis that stochastic resonance is a useful mechanism for low cost information transmission. In any case, however, one would expect that the dependency of the optimal noise level on the input distribution and its dependency on the properties of the neuron population change as soon as metabolic constraints are added.

In our model study we, therefore, explore the complex relationship between information transmission, energy consumption, noise, population size, and the statistics of the input distribution. We do this for three classes of neuron models: binary threshold neuron [17,18], integrate-and-fire neurons [19], and conductance-based point neurons [20]. Binary threshold models are simple enough to be analytically tractable and—together with the integrate-and-fire model—have been widely studied in the context of noisy information transmission. This allows us to directly compare the results of our study with results already published in the literature. Conductance-based point neurons (i.e., Hodgkin-Huxley neuron), on the other hand are biologically more realistic, because the input-, noise-, and activity-induced changes in membrane conductances and their influence on the neuron's dynamics are taken into account. We will, however, see in the following, that the numerical results are (qualitatively) consistent across the different models.

This paper is organized as follows. In Sec. II we investigate the binary threshold model and derive an analytical expression for the optimal noise level. We use an approximation introduced by Brunel and Nadal [21], which is valid in the limit of a large number of neurons. We find that the number of threshold elements for which the analytical expression holds, coincides with the typical number of neurons within a cortical subpopulation of neurons [10]. Compared to other studies (e.g., Ref. [22]), our analytic expression for the optimal noise level does not only depend on the mean of the stimulus distribution, but also on higher moments. Our main results are then verified in Sec. III with the biologically more realistic integrate-and-fire and Hodgkin-Huxley-type models of neurons. We assume that the noise inputs are balanced, i.e., they consist of inhibitory and excitatory inputs with equal efficacy on average. The concept of balanced inputs is biologically plausible and is thought to be a potential mechanism for gain control [23] or rapid state switching in recurrent networks [2]. We show that for populations of spiking neurons, which have a biologically reasonable size, the optimal noise level depends only weakly on the number of neurons. For small populations, however, the noise level has to

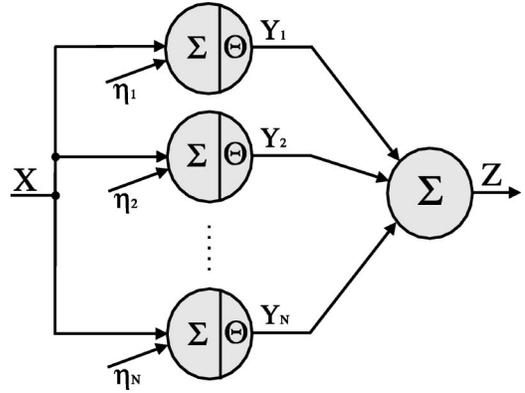


FIG. 1. A population of binary threshold neurons. Each model receives the same input signal and independent Gaussian noise inputs η_i , $i \in \{1, \dots, N\}$ of zero mean and equal variance. Each noise input is independent of the signal and the other noise sources. The total output Z is the sum over all individual outputs Y_i (called pooling in neurophysiological terminology) and is equal to the number of active neurons.

be accurately adjusted to the number of neurons. In Sec. IV, we include metabolic constraints and find that the dependency of the optimal noise level on population size changes dramatically. If information transmission is normalized by metabolic cost as in Refs. [13,15,16], the optimal noise level is almost constant with population size, even for populations with only a few (>5) neurons. When we calculated the input distribution which is optimal with respect to bits per unit metabolic costs, we find that most of the inputs are actually subthreshold.

At first glance, passing information through neurons seems of little use for real neural information processing. An analysis of plain information transmission, however, is a prerequisite for understanding how noise may influence—and improve—transmission of information, after neural computation has been accomplished by dendritic integration. However, some studies [24,25] indicate that optimal information transmission itself can play a major role in natural neural systems. There it has been suggested that a major task of spiny neurons in the barrel field in layer 4 of rat somatosensory cortex is to amplify the weak thalamic input in order to transmit it to the different regions of the cortical column, a scenario, to which the following considerations directly apply.

II. THE BINARY THRESHOLD MODEL

A. Architecture of the model

In this section we consider a population of N binary threshold elements. The total input to each neuron i is the sum of a common input signal X and an individual noise input η_i , and the output Y_i of all these neurons is summed (Fig. 1, see also Ref. [17]).

The output Y_i of the single elements is set to 1 (active), if the total input exceeds a threshold Θ , i.e.,

$$y_i = \begin{cases} 0 & \text{if } x + \eta_i \leq \Theta \\ 1 & \text{if } x + \eta_i > \Theta. \end{cases} \quad (1)$$

In Sec. II, capital letters represent random variables and lower case letters denote the particular realization of the corresponding random variable. The signal X is the same for all threshold elements and is drawn from a Gaussian distribution P_X with mean μ_X and variance σ_X^2 . The noise inputs η_i have a Gaussian distribution with zero mean and variance σ_η^2 , and are mutually independent of the signal X and the other noise sources. Let Z represent the number n of neurons that are set to 1 for a given realization of X . The distribution of Z is then equal to $P_Z(n) = \int_{-\infty}^{\infty} dx P_Z(n|x) P_X(x)$, where the conditional probability $P_Z(n|x)$ can be calculated from

$$P_Z(n|x) = \binom{N}{n} P_{1|x}^n (1 - P_{1|x})^{N-n}. \quad (2)$$

$P_{1|x}$ is the conditional probability that the output of a neuron is set to 1 and is given by

$$P_{1|x} = \int_{\Theta-x}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_\eta} \exp\left(-\frac{\eta^2}{2\sigma_\eta^2}\right) d\eta. \quad (3)$$

B. Approximation of the mutual information

The mutual information is an information theoretic measure [26] which can be used to quantify the amount of information the output Z of the neural population contains about the input X . The mutual information I^{MI} between the input distribution P_X of the signal and the output distribution P_Z is given by

$$I^{MI} = H(Z) - H(Z|X) = - \sum_{n=0}^N P_Z(n) \log_2 P_Z(n) + \int_{-\infty}^{\infty} dx P_X(x) \sum_{n=0}^N P_Z(n|x) \log_2 P_Z(n|x). \quad (4)$$

Brunel and Nadal [21] have shown that in the limit of a large number of neurons, the mutual information between input and output becomes equal to the mutual information between the input signal and an efficient Gaussian estimator of the input signal calculated from the population output. Let

$$\hat{x} = g^{-1}\left(\frac{n}{N}\right) \quad (5)$$

be the maximum likelihood estimator of the signal input x calculated from $P(Z=n|x)$ where $g(\cdot)$ is the error function. This estimator is asymptotically unbiased, efficient, and Gaussian distributed around its mean value. Its variance is $1/F(x)$, where $F(x)$ is the Fisher information:

$$F(x) = E\left[-\frac{\partial^2 \log_2 P(Z=n|x)}{\partial x^2}\right]_x = \left(\frac{\partial P_{1|x}}{\partial x}\right)^2 \frac{N}{P_{1|x}(1-P_{1|x})}. \quad (6)$$

The amount of information the maximum likelihood estimator \hat{X} contains about the stimulus is then given by

$$I(X, \hat{X}) = H(\hat{X}) - \int_{-\infty}^{\infty} dx P_X(x) H(\hat{X}|X=x). \quad (7)$$

In the limit of large N , we can approximate the entropy $H(\hat{X}|X=x)$ with the entropy of a Gaussian distribution with variance $1/F(x)$ and—because the estimator \hat{X} is unbiased—we can replace the entropy $H(\hat{X})$ with $H(X)$. We obtain

$$I(\hat{X}, X) \rightarrow I^F = H(X) - \int_{-\infty}^{\infty} dx P_X(x) \frac{1}{2} \log_2 \left(\frac{2\pi e}{F(x)}\right). \quad (8)$$

Since any processing of signals cannot increase the information content, the mutual information between the signal input and the output of the population is at least equal to I^F [21],

$$I^{MI}(X, Z) \geq I(\hat{X}, X) \geq I^F. \quad (9)$$

We will use I^F to obtain an analytical expression for the optimal noise level σ_{opt}^F . From the condition

$$\frac{\partial I^F}{\partial \sigma_\eta} = -\frac{\partial}{\partial \sigma_\eta} \int_{-\infty}^{\infty} dx P_X(x) \frac{1}{2} \log_2 \left(\frac{2\pi e}{F(x)}\right) = 0, \quad (10)$$

we obtain

$$\sigma_{opt}^F \approx \sqrt{\left(1 - \frac{2}{\pi}\right) (\mu_x^2 + \sigma_x^2)}, \quad (11)$$

where $\log_2[2\pi e/F(x)]$ has been replaced by its second-order Taylor expansion around $x_0=0$ (see Appendix A). In this approximation the optimal noise level depends on the first and second moments of the input distribution, and it is independent of the number of neurons in the population. For non-Gaussian input distributions, higher moments of the input distribution, can be considered if the Taylor expansion is carried out to higher order (see Appendix A).

C. Analysis of information transmission

In Fig. 2(a) the mutual information I^{MI} and its approximation I^F are plotted against the relative strength $\sigma = \sigma_\eta/\sigma_x$ of the noise for various numbers N of neurons in the population. The threshold Θ was set to $\Theta = \mu_x = 0$. The curves show that for increasing size N , I^F becomes a good approximation of the mutual information I^{MI} .

Note that I^F can be negative for small values of noise, because the approximation, Eq. (8), is valid only in the case $F(x) \gg 1$, when the distribution of the estimator \hat{X} is sharply peaked around its mean value. Furthermore, Fig. 2(a) shows that the location σ_{opt}^F of the maximum of the Fisher information is independent of the number of threshold elements, and approaches the optimum of the mutual information I^{MI} as the number of threshold elements increases. For finite N and using Eq. (8), the optimal noise level is overestimated, as can be seen in Fig. 2(b). If the number of neurons in the population is sufficiently large ($N \geq 100$), however, the optimal noise level calculated by I^F , leads to almost optimal infor-

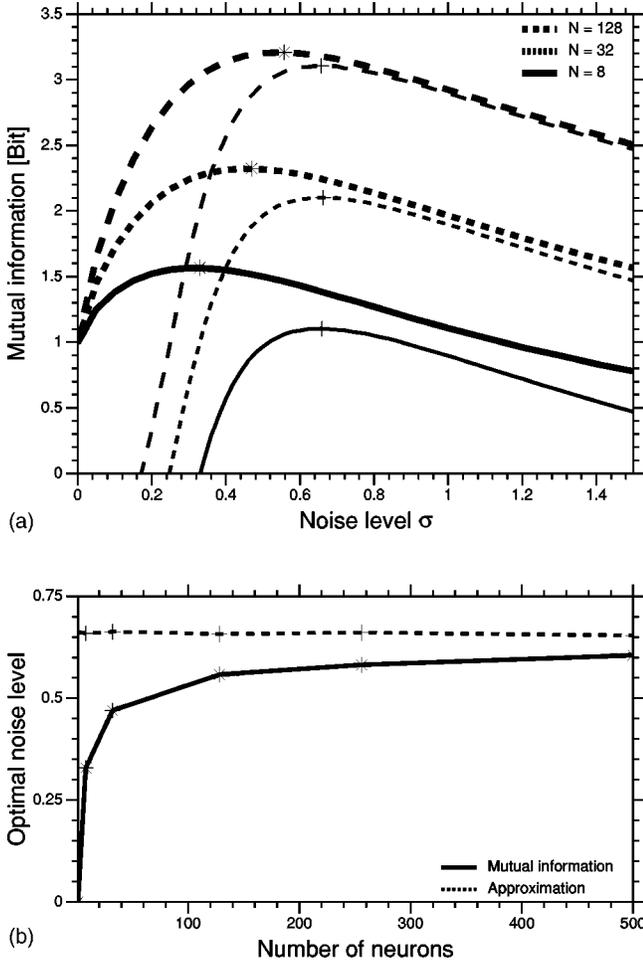


FIG. 2. (a) Mutual information I^{MI} (thick lines) and its approximation I^F according to Eq. (8) (thin lines) as a function of $\sigma = \sigma_\eta/\sigma_x$ for various N (see caption) and for $\Theta = \mu_x = 0$. The mutual information was calculated from Eq. (5), where x was discretized into $100N$ bins. We used $10^5 N$ samples to estimate the distributions $P(Z=n|x)$, $P_Z(n)$, and $P_X(x)$. (b) Optimal noise level of the mutual information (solid line) in comparison with the optimal noise level of I^F obtained from Eq. (8) (dashed line) as a function of the number N of neurons in the population.

mation transmission [see Fig. 2(b)]. Figure 3(a) shows the relative deviation $[|I_{max}^{MI} - I^{MI}(\sigma_{opt}^F)|]/I_{max}^{MI}$ of $I^{MI}(\sigma_{opt}^F)$ from the maximum I_{max}^{MI} of the mutual information as a function of the population size N . $I^{MI}(\sigma_{opt}^F)$ is the mutual information at the optimal noise level of I^F , where σ_{opt}^F was obtained from a Monte Carlo integration of Eq. (8). The figure indicates that the relative deviation is less than 5% for populations with a size of 100 or more neurons. In Fig. 3(b), σ_{opt}^F calculated from Eq. (11), is compared to the optimal noise level of I^F , obtained from Eq. (8). σ_{opt}^F is close to but slightly below the optimal noise level of I^F . Note that σ_{opt}^F is not zero for suprathreshold input signals ($\mu_x > \Theta$), rather information transmission is symmetric with respect to the deviation of μ_x from the threshold θ . This is due to the intrinsic symmetry of the model. Without noise, the output Z is equal to N for all $x > \Theta$; hence noise is needed to distinguish between suprathreshold inputs as well.

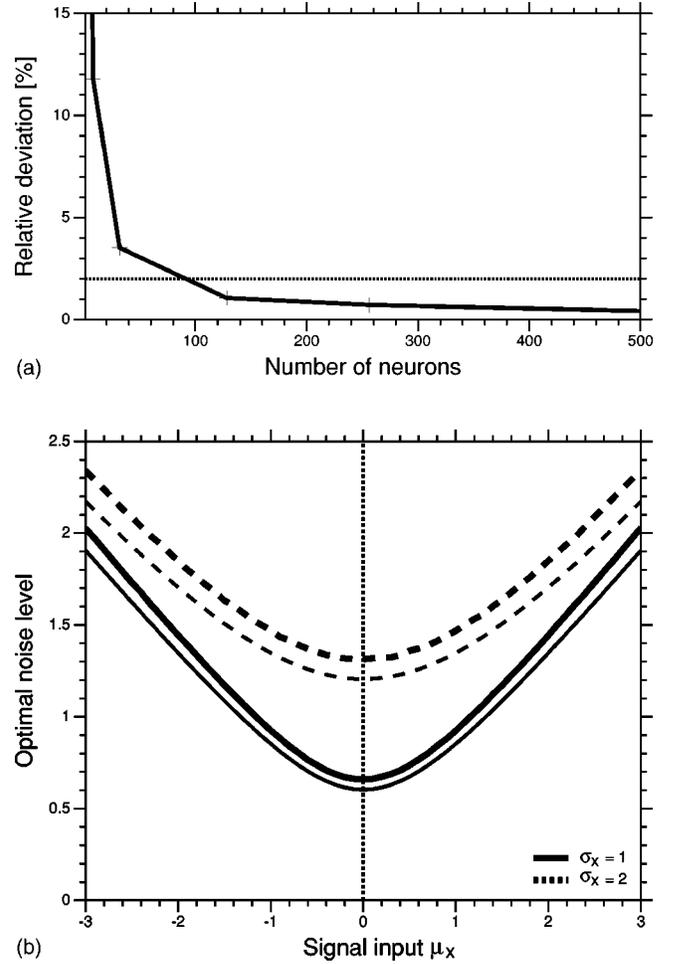


FIG. 3. (a) Relative deviation $[|I_{max}^{MI} - I^{MI}(\sigma_{opt}^F)|]/I_{max}^{MI}$ of $I^{MI}(\sigma_{opt}^F)$ from the maximum I_{max}^{MI} of the mutual information as a function of the population size N . $I^{MI}(\sigma_{opt}^F)$ is the mutual information at the optimal noise level of I^F . The dotted line indicates the 2% level. (b) The optimal noise level according to Eq. (11) (thin line) and according to the maximum of I^F directly obtained from Eq. (8) (thick line) for various values of μ_x . $\sigma_x = 1$ (solid lines) and $\sigma_x = 2$ (dashed lines), $\Theta = 0$.

III. POPULATIONS OF SPIKING NEURONS

To verify the results of the preceding section in a biologically more realistic framework, we replaced the population of binary neurons by a population of spiking neurons, and we chose the leaky integrate-and-fire as well as the Hodgkin-Huxley-type models.

A. Leaky integrate-and-fire neurons

The membrane potential V of the leaky integrate-and-fire neuron changes in time according to the differential equation

$$C_m \frac{dV(t)}{dt} = -g_L[V(t) - E_L] + I_{stim}(t) + \sigma \frac{dW(t)}{dt}, \quad (12)$$

where C_m is the membrane capacitance, g_L the leak conductance of the membrane, E_L the reversal potential, $I_{stim}(t)$ the

external signal, and $dW(t)$ is the infinitesimal increment of a Wiener process which we use to take into account the effect of the noise inputs [19]. Equation (12) describes the sub-threshold dynamics of the membrane potential, as it ignores all active membrane conductances, under the assumption that the synaptic current generated by random synaptic inputs (the background activity) can be approximated by a Wiener process. σ is chosen to be identical for all neurons within the population. Once the membrane potential reaches the threshold, a spike is generated and the membrane potential is clamped to a reset value V_{reset} for an absolute refractory period T_{ref} . A continuous aperiodic Gaussian input signal $I_{stim}(t)$ is generated by a Fourier transform of a band-limited white noise power spectrum, which is added to a constant bias current I_{bias} . For a constant input current below $I_{stim} = 0.5$ nA, the average membrane potential of the neurons remains subthreshold. The parameters used for the simulations are given in Appendix B.

B. Hodgkin-Huxley neurons

The leaky integrate-and-fire neuron is widely used as a building block of neural network models, because of its simplicity. However, this model does not account for two properties of real neurons, which may turn out to be important in the context of this modeling study. First, it does not include the changes in the membrane conductance caused by the synaptic input. Second, the membrane potential is reset to a certain fixed value after each spike, an assumption which may not capture the actual changes in the membrane potential. The simplest model which accounts for the above mentioned phenomena is a single-compartment Hodgkin-Huxley model. Models of this kind are highly successful in describing experimental data (for an introduction, see Ref. [27]). The membrane potential V of the Hodgkin-Huxley neurons we used in this study changes in time according to the differential equation

$$C_m \frac{\partial V}{\partial t} = -g_L(V - E_L) - I_{Na} - I_K - I_M - I_{syn} + I_{stim}(t). \quad (13)$$

The left-hand side of the equation describes the influence of the membrane's capacitance, while all ionic currents through the cell's membrane—including the noise and stimulus currents—are summed on the right-hand side. We consider a leak current $I_L = g_L(V - E_L)$, and the spike-generating sodium $[I_{Na} = \bar{g}_{Na} m^3 h (V - E_{Na})]$ and potassium $[I_K = \bar{g}_K n^4 (V - E_K)]$ currents as well as an additional noninactivating potassium current $I_M = \bar{g}_M p (V - E_K)$, which is responsible for spike frequency adaptation and which is typically found in neocortical pyramidal cells [20]. The parameters used for the simulations are given in Appendix C. The total synaptic noise current I_{syn} is generated by the changing membrane conductance induced by stochastic spike trains arriving at excitatory and inhibitory synapses. The aperiodic Gaussian stimulus $I_{stim}(t)$ is generated as described in the preceding section. A detailed description of the currents can be found in Ref. [1].

To obtain a more realistic approximation of the synaptic background activity, we used a point-conductance model recently described by Destexhe *et al.* [20]. In this model the total synaptic current I_{syn} is obtained as a superposition of excitatory (e) and inhibitory (i) inputs,

$$I_{syn} = g_e(t)(V - E_e) + g_i(t)(V - E_i), \quad (14)$$

where g_e and g_i are the synaptic conductances, and E_e and E_i the corresponding reversal potentials. The time-dependent conductances are described as an Ornstein-Uhlenbeck process,

$$\frac{dg_e(t)}{dt} = -\frac{1}{\tau_e} [g_e(t) - g_{e0}] + \sigma_{e0} \frac{dW(t)}{dt}, \quad (15)$$

$$\frac{dg_i(t)}{dt} = -\frac{1}{\tau_i} [g_i(t) - g_{i0}] + \sigma_{i0} \frac{dW(t)}{dt}. \quad (16)$$

The parameter of the Ornstein-Uhlenbeck processes are chosen in such a way that they resemble *in vivo* like activity [20] (see Appendix D). The average synaptic noise current is close to zero (i.e., balanced) just below threshold, because excitation and inhibition cancel. We used balanced noise input, because it leads to an increase of response variability without changing the average membrane potential and because it is a plausible model for the synaptic input to cortical neurons introduced by background activity [23]. However, an exact balance between excitation and inhibition is only possible for a given fixed membrane voltage V . For other values, a change in the strength of the noise input also leads to a shift of the average potential. Here we choose a ratio of $g_{i0}/g_{e0} = 3.2$, for which balance is achieved at approximately 3.5 mV below threshold. Different noise conditions are modeled by multiplying the synaptic conductances $[(g_{e0}, g_{i0})]$ and the standard deviations $[(\sigma_{e0}, \sigma_{i0})]$ by a common gain factor α .

C. Quantification of information transmission

Information transmission through a population of spiking neurons is much harder to estimate than information transfer through binary threshold elements, because the amount of data needed to get a reasonable estimate of the probability distributions explodes. Here we use an approach which was recently described in literature [27,28]. The main idea of this approach is that a lower bound on the information rate can be obtained by computing an estimate of the input signal from the observed spike train. The estimate is calculated with a method called Wiener-Kolmogorov filtering, and contains no information, which was not actually present in the spike train. To get the estimate, the spike train $z(t)$ is convolved with a filter h , which minimizes the mean square error,

$$\epsilon^2(h) = \langle |I_{stim}(t) - h * z(t)|^2 \rangle, \quad (17)$$

between the stimulus $I_{stim}(t)$ and its estimate $I_{est}(t) = h * z(t)$. This filter h represents a noncausal optimal linear filter of the spike train and can be obtained by solving the condition $d\epsilon^2(h)/dh = 0$ for h . This yields $h(w)$

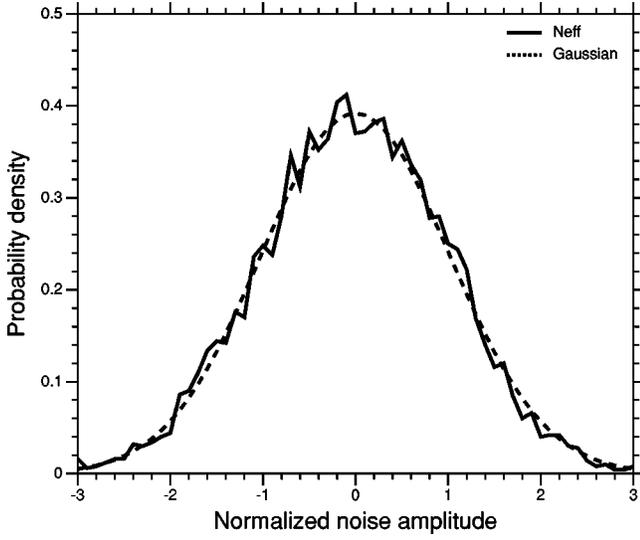


FIG. 4. Distribution of the normalized effective noise level N_{eff} in comparison to a Gaussian distribution with the same variance. The histogram is constructed from 5000 segments of 819 ms duration ($dt=0.2$ ms). The bias current was set to $I_{bias}=0.5$ nA, the standard deviation of the stimulus was set to $std(I_{stim})=0.2$ nA, and the input noise level was equal to $\sigma=1$. Numerical simulations for different levels σ of input noise and different values of I_{bias} lead to similar results.

$=S_{IX}(-\omega)/S_{XX}(\omega)$, where $S_{XX}(\omega)$ is the power spectrum of the spike train and $S_{IX}(\omega)$ denotes the Fourier transform of the cross correlation of the spike train and the stimulus. Next we calculated the Fourier components of the effective noise $n_{eff}(\omega)$, via the relation

$$I_{est}(\omega) = g(\omega)[I_{stim}(\omega) - n_{eff}(\omega)], \quad (18)$$

where $I_{stim}(\omega)$ and $I_{est}(\omega)$ are the Fourier transforms of the input signal and its estimate, $g(\omega)$ is the frequency-dependent gain introduced to correct for systematic errors, which arise from filtering the spike train (see Rieke *et al.* [28] for a further information), and n_{eff} is the effective noise in the estimate. The stimulus and its estimate are divided into segments of approximately 1 sec, and the Fourier transform of each segment is calculated. Given enough segments, the frequency-dependent gain $g(\omega)$ and the Fourier components of the effective noise $n_{eff}(\omega)$ can be determined using linear regression [cf. Eq. (18)]. The power spectrum of the effective noise level $N_{eff}(\omega)$ is obtained by calculating the variance of the noise components $n_{eff}(\omega)$ normalized by the time window T_n , $N_{eff}(\omega) = \text{var}(n_{eff}(\omega))T_n$. Finally, we calculate the lower bound R_{info} to the information rate using

$$R_{info} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \log_2 \left[1 + \frac{S(\omega)}{N_{eff}(\omega)} \right], \quad (19)$$

which is close to the real information rate, if the effective noise is approximately Gaussian distributed [28]. This is the case here as shown in Fig. 4.

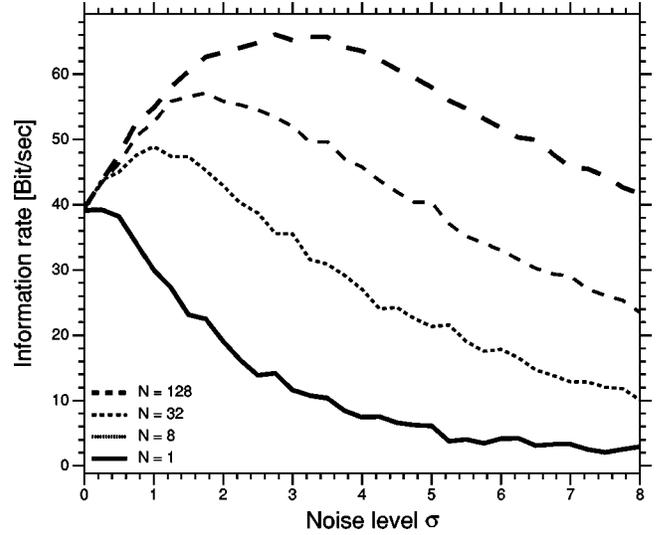


FIG. 5. The leaky integrate-and-fire neuron. The information rate calculated from Eq. (19) is plotted as a function of the input noise for different numbers of neurons in the population. The bias current was set to $I_{bias}=0.5$ nA. The standard deviation of the stimulus was equal to $std(I_{stim})=0.2$ nA and the information rate was calculated from 200 samples of 819 ms duration ($dt=0.2$ ms). The threshold is located at 0.5 nA.

D. Results of numerical simulations

Figure 5 shows the results obtained with the leaky integrate-and-fire model. The lower bound R_{info} of the information rate is plotted as a function of the input noise for different numbers of neurons in the population. The bias current was set to $I_{bias}=0.5$ nA and the standard deviation was equal to $std(I_{stim})=0.2$ nA. Figure 5 shows that the optimal noise level increases with the increase in the number of neurons in the population, similar to the binary threshold model. The optimal noise level is plotted as a function of the population size in Fig. 7(a) (dashed line). It depends weakly (close to logarithmically) on the number of neurons in a biologically reasonable range.

Figure 6 shows the results obtained with the Hodgkin-Huxley model. R_{info} is plotted as a function of the noise level for different number of neurons. Because the mean firing rate of the Hodgkin-Huxley neurons is smaller than the mean firing rate of the leaky integrate-and-fire neurons for the same stimulus, the information rate is smaller than the rate given in Fig. 5. Again, the optimal noise level [Fig. 7(a), solid line] depends weakly (close to logarithmically) on the number of neurons. Note, that the maxima of the information rate are broad in both cases (Figs. 5 and 6), i.e., the amount of transmitted information degrades only slightly if the level of input noise deviates from its optimal value.

Figure 7(b) shows, how information transmission depends on the number of neurons in the population, for a given level of noise. The noise level was set to its optimal value for a population size of $N=100$ neurons. The solid line shows the results for the integrate-and-fire model. The relative deviation $(|I_{max}^{IR} - I^R|)/I_{max}^{IR}$ from the maximum I_{max}^{IR} of the information rate is plotted for different numbers of neurons. Figure 7(b) shows that the loss in information transmission is

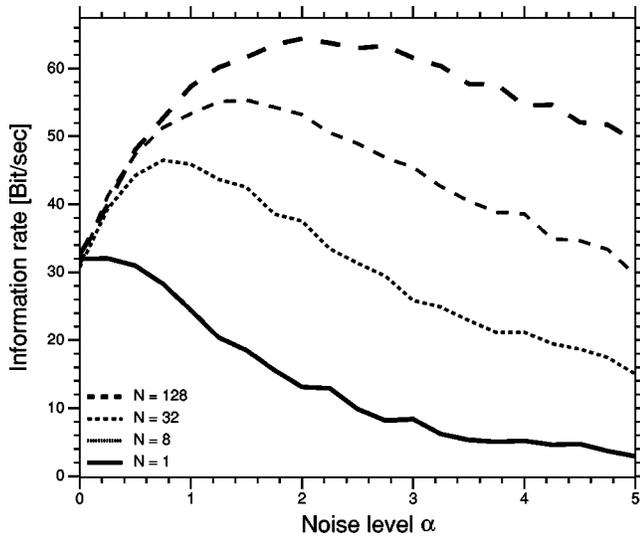


FIG. 6. The Hodgkin-Huxley neuron. The information rate calculated from Eq. (19) is plotted as a function of the input noise for different numbers of neurons in the population. The bias current was set to $I_{bias}=0.5$ nA. The standard deviation of the stimulus was equal to $\text{std}(I_{stim})=0.2$ nA, and the information rate was calculated from 200 samples of 819 ms duration ($dt=0.2$ ms). The threshold is located at 0.5 nA.

less than 2% for populations of neurons with a population size between 80 and 130 neurons. The dashed line in Fig. 7(b) shows the corresponding results for the Hodgkin-Huxley model. Again, the loss in information transmission is small.

IV. ENERGY EFFICIENT INFORMATION TRANSMISSION

Information transmission in the brain is metabolically expensive [14]. Especially, the generation of spikes consumes a huge amount of energy. If the cost of firing is high in comparison to the “housekeeping” cost, than it is advantageous for the brain to use energy efficient neural codes [13–16]. In such a case one would expect that the dependency of the optimal noise level on the population size will also change. We therefore investigate the role of noise for energy efficient information transmission in the following two chapters.

Given a fixed amount of energy, there are several ways to achieve energy efficient information transmission. One strategy is to use an input distribution so that the energy constraint can be fulfilled. Another strategy is to maximize the information rate per metabolic cost to transmit as much information as possible. In the following, we will take a closer look on both strategies.

A. Optimal input distribution

In the previous chapter we showed that the maximum of the information transmission depends on the number of neurons in the population and on the input distribution. If the network is forced to use less energy for transmission, than the optimal input distribution is shifted into the subthreshold regime, as we will show in the following for the binary threshold model.

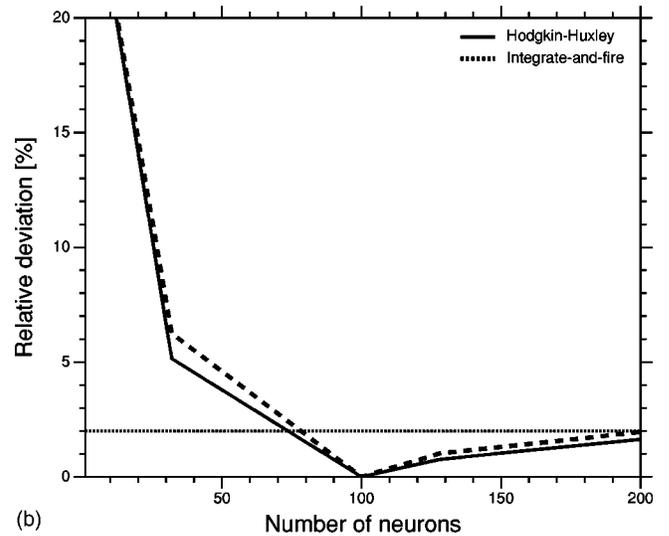
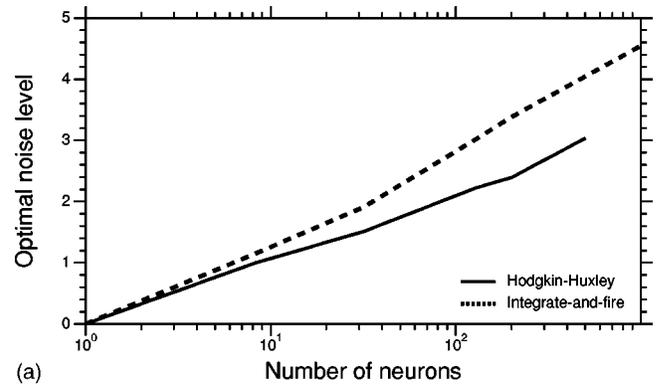


FIG. 7. (a) The optimal noise level as a function of the number of neurons in the population for the integrate-and-fire model (dashed line) in comparison with the Hodgkin-Huxley model (solid line). (b) Relative deviation from the maximal information rate as a function of the number of neurons in the array for the integrate-and-fire model (dashed line) in comparison to the Hodgkin-Huxley model (solid line). The noise level was chosen to be optimal for a population size $N=100$. Parameters for (a,b) as in Figs. 5 and 6.

We calculate the probability distribution $P_X(x)$ of the input signal, which maximizes the information transmission under the constraint that the average cost of transmission is kept below a given energy E_{max} ,

$$\max_{P(x)} I(X,Z) \quad \text{with} \quad \bar{E} = \sum_x E(x)P(x) < E_{max}. \quad (20)$$

In order to solve this optimization problem, we use an iterative algorithm which is based on the work of Arimoto [29] and Blahut [30], and which was used in several other studies about energy efficient coding [15,16,31]. The input distribution $P_X(x)$ was discretized and the cost of transmission $E(x)$ for each x was defined to be equal to the average number of neurons set to 1, i.e., $E(x) = b + \sum_n nP(Z=n|x)$, where $P(Z=n|x)$ is defined in Eq. (2) and b is the fixed baseline cost which can be chosen arbitrary. Because the

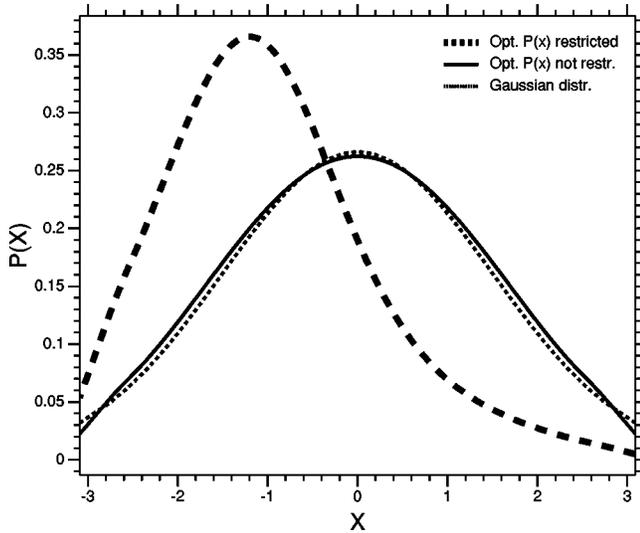


FIG. 8. Optimal input distribution for the binary threshold model. The solid line shows the input distribution which maximizes the information transmission through the network measured by the mutual information $I(X,Z)$ and without metabolic constraints. The distribution is close to a Gaussian distribution (dotted line). The average metabolic costs in this case are $E_{max}=(N/2)+b$. The dashed line shows the optimal input distribution under the metabolic constraint $E_{max}\leq(N/4)+b$. The distribution is shifted to the subthreshold regime. The optimal input distribution was calculated for the variable $P_{1|x}$ (discretized to a resolution $\Delta P=0.002$) and was transformed back to the input space ($N=10\,000$, $\Theta=0$, and $\sigma_\eta=1$).

range of the input X lies between ∞ and $-\infty$, we calculated the optimal input distribution for the transformed variable $P_{1|x}$ (lying in the range $[0,1]$, discretized to a resolution $\Delta P=0.002$) and transformed it back to the input space. The number of neurons in the population was set to $N=10\,000$. Figure 8 shows the optimal input distributions calculated with the Blahut-Arimoto algorithm. The solid line is the optimal input distribution for unlimited energy consumption. In this case the average energy consumption is equal to $E_{max}=(N/2)+b$. As one can see in Fig. 8, the optimal input distribution is symmetric around the threshold $\Theta=0$ and very close to a Gaussian distribution (dotted line). Limiting the amount of metabolic costs [$E_{max}=(N/4)+b$] destroys the symmetry and leads to an input distribution which is mainly subthreshold (dashed line). How much the optimal input distribution is shifted into the subthreshold regime depends on the value E_{max} , but clearly, energy efficient codes favor low rates and subthreshold inputs.

B. Information rate per metabolic cost

As a measure of metabolic efficiency, we now consider the ratio between the transmitted information R_{info} and the total metabolic cost E . We assume that the average metabolic cost E per unit time is a sum of a term proportional to the average rate \bar{r} of the neurons and a term which contributes a fixed baseline cost b . The baseline cost represents the metabolic expense of maintaining a single neuron in the population. We set

$$E=cN(b+\bar{r}), \quad (21)$$

where N is the number of neurons and c is a proportionality constant, which can be interpreted as the average cost per spike. Both information rate and average cost change with the actual noise level. We assume that the average cost is proportional to the number of neurons in the population as in Refs. [13,14]. To investigate the relationship between the noise level and the information rate per unit cost, we vary the noise level in the population of neurons for a given population size N .

Figure 9 shows the results obtained from the Hodgkin-Huxley model. In Fig. 9(a) the optimal noise level for the unconstrained case (dashed line) is compared with the optimal noise level of the information rate per unit cost for different baseline cost ($b=0$ thick solid line, $b=5$ thick dotted line, $b=20$ thin solid line, $b=100$ thin dotted line) for a subthreshold input. If metabolic costs are taken into account, the dependency of the optimal noise level on the population size N almost vanishes even for a small population size. This is due to the fact that high energy costs favor low output rates, which—in turn—can be achieved if the level of noise is reduced. If baseline costs increase, so does the optimal noise level, because the smaller dependency of the total cost on the output activity \bar{r} allows for higher output rates. In the limit of large baseline costs, the activity dependent costs ($\bar{r}N$) can be neglected, and the optimal noise level of the information rate per unit cost is equal to the optimal noise level for the unconstrained case.

For less dominating baseline cost, the information transmission becomes more and more robust against changes in the population size even for small populations. Note that the deviation from the maximum (information rate per cost) is less than 2% for populations of neurons with 3 to at least 50 neurons [Fig. 9(b)]. For example, if the neurons in the population adapt to the optimal noise level of a population of seven neurons, the relative deviation from the maximum is less than 2% for populations of neurons with population size between 5 and 21 neurons [see Fig. 9(b)].

V. DISCUSSION

In this model study we examined how background activity affects the information transmission in a population of neurons. Using an abstract framework based on binary threshold elements we showed that the optimal noise level of the mutual information I^{MI} between an input distribution and the corresponding output can be approximated by the Fisher information I^F , calculated from Eq. (8), for large enough population size, and that the quality of this approximation increases with the number N of neurons in the population. In general, the optimal noise level σ_{opt}^{MI} for I^{MI} depends on N , thus a neuron that should adapt to the optimal noise level must have some knowledge about the size of the population. We showed that for all N , σ_{opt}^F is larger than σ_{opt}^{MI} , and that with increasing N this difference decreases. Due to the broad maxima and the asymmetry of the stochastic resonance curves (transmitted information vs noise), a moderate over-

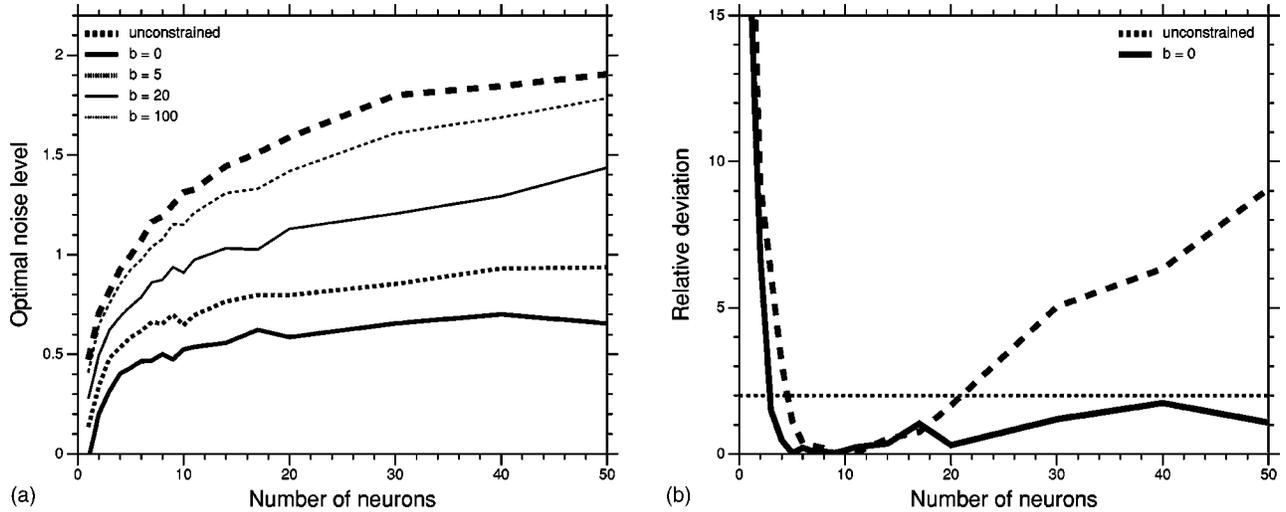


FIG. 9. (a) Optimal noise level of the information rate per cost plotted against the number of Hodgkin-Huxley neurons in the population for different baseline cost ($b=0$: thick solid line, $b=5$: thick dotted line, $b=20$: thin solid line, and $b=100$: thin dotted line). The dashed line is the optimal noise level for the unconstrained case. (b) Relative deviation from the maximum of the information rate per cost as a function of the number of Hodgkin-Huxley neurons in the population for different baseline costs ($b=0$: solid line, unconstrained case: dashed line). The noise level was chosen to be optimal for a population size $N=7$. Parameters for (a) and (b) are: $I_{bias}=0.35$ nA, $\text{std}(I_{stim})=0.2$ nA, and $dt=0.2$ ms.

estimation of the optimal noise level does not degrade the amount of information transmitted in a dramatic way, provided that the number of neurons in the population is sufficiently large. Thus, adaptation to σ_{opt}^F instead of σ_{opt}^{MI} is reasonable for population sizes $N \geq 100$. Furthermore, the analytic expression for the optimal noise level, in terms of I^F , does not only depend on the mean of the input distribution, as in Ref. [22], but also on the variance. More generally, evaluating the integral in Eq. (8) using more terms in the Taylor expansion yields a weighted sum of the moments of the input distribution.

In the more realistic leaky-integrate-and-fire and the Hodgkin-Huxley framework, respectively, the optimal noise level also depends strongly on N for small populations, but the dependency becomes weak, if N is large enough. These results have consequences for a possible adaptation of the neurons' noise input to changing distribution of signal inputs. If N is large adaptation is local in the sense that it does require only those quantities which are locally available at the single neuron. Note that this would be the case for the size of typical cell assemblies in cortex which has been estimated to be $N \approx 200$ [10]. If N is small, the background activity can have a wide influence of the information transmission properties of the population and therefore should be adjusted accurately in each single neuron.

Recent studies have shown that neural systems prefer information transmission via many parallel low intensity channels [14,15]. Similar to that, our simulations of the binary threshold model have shown that limiting the usable energy for transmission leads to a shift of the optimal input distribution to the subthreshold regime. Since application of noise is one way to allow for transmission of otherwise subthreshold signals, the strive for energy efficient codes may be a justification of stochastic resonance in neural systems. Taking the cost of information transmission into account, the

dependency of the optimal noise level of the ratio between information rate and average cost on the number of neurons N becomes very weak. This holds even for small populations, $N < 10$, provided that the baseline costs are small compared to the rate dependent costs. Thus, noise can contribute to the enhancement of information transmission, and may be adapted via a learning rule, which depends on single-neuron properties only, even when the number of neurons in the population is small.

ACKNOWLEDGMENTS

The authors thank A. R. Garg and L. Schwabe for fruitful discussions. The graphics were created with the epsTk-toolkit written by Stefan Müller. This work was supported by the DFG (Grant No. SFB 618).

APPENDIX A: OPTIMIZATION OF I_F WITH RESPECT TO THE NOISE LEVEL

The Fisher information I_F is given by

$$I_F = H(X) - \int_{-\infty}^{\infty} dx P_X(x) \frac{1}{2} \log_2 \left(\frac{2\pi e}{F(x)} \right). \quad (\text{A1})$$

Because the entropy of the input distribution $H(X)$ does not depend on the noise level, it is sufficient to minimize

$$\int_{-\infty}^{\infty} dx P_X(x) \ln \left(\frac{1}{F(x)} \right) = \min. \quad (\text{A2})$$

Inserting Eq. (6) for $F(x)$ we get

$$\int_{-\infty}^{\infty} dx P_X(x) \left(-2 \ln \frac{\partial P_{1|x}}{\partial x} - \ln N + \ln [P_{1|x}(1 - P_{1|x})] \right) = \min. \quad (\text{A3})$$

Under the assumption of independent Gaussian noise we obtain for $\partial P_{1|x}/\partial x$ [see Eq. (3)],

$$\frac{\partial P_{1|x}}{\partial x} = \frac{1}{\sqrt{2\pi}\sigma_\eta} \exp\left(-\frac{(\Theta-x)^2}{2\sigma_\eta^2}\right). \quad (\text{A4})$$

We now replace $\ln[P_{1|x}(1 - P_{1|x})]$ by its second-order Taylor expansion around $x = \Theta$,

$$\ln[P_{1|x}(1 - P_{1|x})] = -\ln 4 - \frac{2}{\pi} \frac{(x - \Theta)^2}{\sigma_\eta^2} + O(x^4) \quad (\text{A5})$$

and obtain

$$\int_{-\infty}^{\infty} dx P_X(x) \left[2 \ln \sigma_\eta + \left(1 - \frac{2}{\pi}\right) \frac{(x - \Theta)^2}{\sigma_\eta^2} \right] = \min, \quad (\text{A6})$$

where all terms independent of σ_η have been omitted. Now we define $\Theta = 0$ as origin and solve the integral which yields

$$2 \ln \sigma_\eta + \left(1 - \frac{2}{\pi}\right) \frac{(\mu_x^2 + \sigma_x^2)}{\sigma_\eta^2} = \min. \quad (\text{A7})$$

Setting the derivative with respect to σ_η to zero, we finally obtain

$$\sigma_{opt}^F = \sqrt{\left(1 - \frac{2}{\pi}\right) (\mu_x^2 + \sigma_x^2)}. \quad (\text{A8})$$

If the Taylor expansion [see Eq. (A5)] is extended to higher-order, then the evaluation of the integral in Eq. (A6) leads to higher order moments of the input distribution. For example, taking the Taylor expansion to fourth order, the evaluation of the integral

$$\int_{-\infty}^{\infty} dx P_X(x) \left[2 \ln \sigma_\eta + \left(1 - \frac{2}{\pi}\right) \frac{(x - \Theta)^2}{\sigma_\eta^2} + \frac{2}{3\pi} \left(1 - \frac{3}{\pi}\right) \frac{(x - \Theta)^4}{\sigma_\eta^4} \right] = \min \quad (\text{A9})$$

yields the following optimal noise level:

$$(\sigma_{opt}^F)^2 = \left(\frac{1}{2} - \frac{1}{\pi}\right) (\mu_x^2 + \sigma_x^2) \pm \frac{\sqrt{3}}{6\pi} \sqrt{a_1 \mu_x^4 + 6a_2 \mu_x^2 \sigma_x^2 + 3a_2 \sigma_x^4}. \quad (\text{A10})$$

Parameters a_1 and a_2 are $a_1 = 3\pi^2 + 4\pi - 36$ and $a_2 = \pi^2 + 12\pi - 44$.

APPENDIX B: THE LEAKY INTEGRATE-AND-FIRE NEURON

Dynamics of the membrane potential is given by

$$C_m \frac{dV(t)}{dt} = -g_L(V(t) - E_L) + I_{stim}(t) + \sigma \frac{dW(t)}{dt}. \quad (\text{B1})$$

Model parameters are $\tau_m = 20$ ms, $g_L = 25$ nS, $C_m = \tau_m g_L = 0.5$ nF, $E_L = -74$ mV, $V_{reset} = -60$ mV, $T_{refrac} = 1.72$ ms, and threshold is -54 mV. Parameters are taken from [32]. $I_{stim}(t)$ is generated by a Fourier transform of a white noise power spectrum with a cutoff frequency of 20 Hz. All simulations were done using the Euler integration scheme with exact update equation [33] and a fixed time step of $dt = 0.2$ ms.

APPENDIX C: THE HODGKIN-HUXLEY NEURON

Dynamics of the membrane potential is given by

$$C_m \frac{\partial V}{\partial t} = -g_L(V - E_L) - I_{Na} - I_K - I_M - I_{syn} + I_{stim}(t). \quad (\text{C1})$$

Model parameters are $C_m = 1$ $\mu\text{F}/\text{cm}^2$, $g_L = 0.045$ mS, and $E_L = -70$ mV.

Voltage-dependent sodium current, I_{Na} , given by

$$I_{Na} = \bar{g}_{Na} m^3 h (V - E_{Na}),$$

$$\frac{dm}{dt} = \alpha_m(V)(1 - m) - \beta_m(V)m,$$

$$\frac{dh}{dt} = \alpha_h(V)(1 - h) - \beta_h(V)h,$$

$$\alpha_m(V) = \frac{-0.32(V - V_T - 13)}{\exp[-(V - V_T - 13)/4] - 1},$$

$$\beta_m(V) = \frac{0.28(V - V_T - 40)}{\exp[-(V - V_T - 40)/5] - 1},$$

$$\alpha_h(V) = 0.128 \exp[-(V - V_T - V_S - 17)/18],$$

$$\beta_h(V) = \frac{4}{1 + \exp[-(V - V_T - V_S - 40)/5]}.$$

Model parameters are $\bar{g}_{Na} = 3$ mS/cm², $V_T = -58$ mV, and $V_S = -10$ mV.

Delayed-rectifier potassium current, I_K given by

$$I_K = \bar{g}_K n^4 (V - E_K),$$

$$\frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n,$$

$$\alpha_n(V) = \frac{-0.032(V - V_T - 15)}{\exp[-(V - V_T - 15)/5] - 1},$$

$$\beta_n(V) = 0.5 \exp[-(V - V_T - 10)/40].$$

Model parameters are $\bar{g}_K = 5 \text{ mS/cm}^2$ and $V_T = -58 \text{ mV}$.
Noninactivating potassium current, I_M given by

$$I_M = \bar{g}_M p (V - E_K),$$

$$\frac{dp}{dt} = \alpha_p(V)(1-p) - \beta_p(V)p,$$

$$\alpha_p(V) = \frac{0.0001(V + 30)}{1 - \exp[-(V + 30)/9]},$$

$$\beta_p(V) = \frac{-0.0001(V + 30)}{1 - \exp[(V + 30)/9]}.$$

Model parameters are $\bar{g}_M = 1 \mu\text{ S/cm}^2$. The spike threshold is $\approx -55 \text{ mV}$. The simulations of the population of Hodgkin-Huxley neurons were done with the NEURON simulation environment [34] using the model from Destexhe *et al.* described above [20].

APPENDIX D: SYNAPTIC BACKGROUND ACTIVITY

Synaptic current given by

$$I_{syn} = g_e(t)(V - E_e) + g_i(t)(V - E_i). \quad (\text{D1})$$

Model of synaptic conductances, given by

$$\frac{dg_e(t)}{dt} = -\frac{1}{\tau_e}[g_e(t) - \alpha g_{e0}] + \alpha \sigma_{e0} \frac{dW(t)}{dt}, \quad (\text{D2})$$

$$\frac{dg_i(t)}{dt} = -\frac{1}{\tau_i}[g_i(t) - \alpha g_{i0}] + \alpha \sigma_{i0} \frac{dW(t)}{dt}. \quad (\text{D3})$$

Model parameters are $g_{e0} = 0.01 \mu\text{ S}$, $g_{i0} = 0.032 \mu\text{ S}$, $\sigma_{e0} = 0.003 \mu\text{ S}$, $\sigma_{i0} = 0.00825 \mu\text{ S}$, $\tau_e = 2.7 \text{ ms}$, $\tau_i = 10.5 \text{ ms}$.

-
- [1] A. Destexhe and D. Paré, *J. Neurophysiol.* **81**, 1531 (1999).
[2] M. Tsodyks and T. Sejnowski, *Network Comput. Neural Syst.* **6**, 111 (1995).
[3] M. Rudolph and A. Destexhe, *J. Comput. Neurosci.* **11**, 19 (2001).
[4] L. Gammitoni, P. Hänggi, P. Jung, and F. Marchesoni, *Rev. Mod. Phys.* **70**, 223 (1998).
[5] A. Longtin, *J. Stat. Phys.* **70**, 309 (1993).
[6] J.K. Douglass, L.A. Wilkens, E. Pantazelou, and F. Moss, *Nature (London)* **365**, 337 (1993).
[7] D.F. Russel, L.A. Wilkens, and F. Moss, *Nature (London)* **402**, 291 (1999).
[8] G. Wenning and K. Obermayer, *Phys. Rev. Lett.* **90**, 120602 (2003).
[9] A. Pouget, R.S. Zemel, and P. Dayan, *Nature Rev. Neurosci.* **1**, 125 (2000).
[10] M.L. Feldman, in *Morphology of the Neocortical Pyramidal Neuron*, edited by A. Peters and E.G. Jones, *Cerebral Cortex Vol. 1* (Plenum, New York, 1984).
[11] J.J. Collins, C.C. Chow, and T.T. Imhoff, *Nature (London)* **376**, 236 (1995).
[12] N.G. Stocks and R. Mannella, *Phys. Rev. E* **64**, 030902 (2001).
[13] W.B. Levy and R.A. Baxter, *Neural Comput.* **8**, 531 (1996).
[14] S.B. Laughlin, R.R. de Ruyter van Steveninck, and J.C. Anderson, *Nat. Neurosci.* **1**, 36 (1998).
[15] V. Balasubramanian, D. Kimber, and M.J. Berry II, *Neural Comput.* **13**, 799 (2001).
[16] G.G. de Polavieja, *J. Theor. Biol.* **214**, 657 (2002).
[17] N.G. Stocks, *Phys. Rev. Lett.* **84**, 2310 (2000).
[18] N.G. Stocks, *Phys. Rev. E* **63**, 041114 (2001).
[19] H.C. Tuckwell, *Introduction to Theoretical Neurobiology* (Cambridge University Press, Cambridge, 1988), Vol. 2.
[20] A. Destexhe, M. Rudolph, J. Fellous, and T.J. Sejnowski, *Neuroscience* **107**, 13 (2001).
[21] N. Brunel and J. Nadal, *Neural Comput.* **10**, 1731 (1998).
[22] M. Stemmler, *Network Comput. Neural Syst.* **7**, 687 (1996).
[23] F.S. Chance, L.F. Abbott, and A.D. Reyes, *Neuron* **35**, 773 (2002).
[24] J. Lübke, V. Egger, B. Sakmann, and D. Feldmeyer, *J. Neurosci.* **20**, 5300 (2000).
[25] D. Feldmeyer, V. Egger, J. Lübke, and B. Sakmann, *J. Physiol.* **521**, 169 (1999).
[26] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, New York, 1991).
[27] *Methods in Neural Modeling*, 2nd ed., edited by C. Koch and I. Segev (MIT Press, Cambridge, 1998).
[28] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, 1997).
[29] S. Arimoto, *IEEE Trans. Inf. Theory* **18**, 14 (1972).
[30] R.E. Blahut, *IEEE Trans. Inf. Theory* **18**, 460473 (1972).
[31] S. Schreiber, C.K. Machens, A.V.M. Herz, and S.B. Laughlin, *Neural Comput.* **14**, 1323 (2002).
[32] E. Salinas and T.J. Sejnowski, *J. Neurosci.* **20**, 6193 (2000).
[33] D. Gillespie, *Phys. Rev. E* **54**, 2084 (1996).
[34] M.L. Hines and N.T. Carnevale, *Neural Comput.* **9**, 1179 (1997).