Energy-Efficient Coding with Discrete Stochastic Events

Susanne Schreiber
susanne@salk.edu
Institute of Biology, Humboldt-University Berlin, 10115 Berlin, Germany, and
Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, U.K.

Christian K. Machens
c.machens@itb.biologie.hu-berlin.de
Andreas V. M. Herz
a.herz@biologie.hu-berlin.de
Institute of Biology, Humboldt-University Berlin, 10115 Berlin, Germany

Simon B. Laughlin
s.laughlin@zoo.cam.ac.uk
Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, U.K.

We investigate the energy efficiency of signaling mechanisms that transfer information by means of discrete stochastic events, such as the opening or closing of an ion channel. Using a simple model for the generation of graded electrical signals by sodium and potassium channels, we find optimum numbers of channels that maximize energy efficiency. The optima depend on several factors: the relative magnitudes of the signaling cost (current flow through channels), the fixed cost of maintaining the system, the reliability of the input, additional sources of noise, and the relative costs of upstream and downstream mechanisms. We also analyze how the statistics of input signals influence energy efficiency. We find that energy-efficient signal ensembles favor a bimodal distribution of channel activations and contain only a very small fraction of large inputs when energy is scarce. We conclude that when energy use is a significant constraint, trade-offs between information transfer and energy can strongly influence the number of signaling molecules and synapses used by neurons and the manner in which these mechanisms represent information.

1 Introduction

Energy and information are intimately related in all forms of signaling. Cellular signaling involves local movements of ions and molecules, shifts in their concentration, and changes in molecular conformation, all of which

require energy. Nervous systems have highly evolved cell signaling mechanisms to gather, process, and transmit information, and the quantities of energy consumed by neural signaling can be significant. In the blowfly retina, the transmission of a single bit of information across one chemical synapse requires the hydrolysis of more than $100,000$ ATP molecules (Laughlin, de Ruyter van Steveninck, & Anderson, 1998). The adult human brain accounts for approximately 20% of resting energy consumption (Rolfe & Brown, 1997). Recent calculations suggest that the high rate of energy consumption in cortical gray matter results mainly from the transmission of electrical signals along axons and across synapses (Attwell & Laughlin, 2001). Given that high levels of energy consumption constrain function, it is advantageous for nervous systems to use energy-efficient neural mechanisms and neural codes (Levy & Baxter, 1996; Baddeley et al., 1997; Sarpeshkar, 1998; Laughlin, Anderson, O'Carroll, & de Ruyter van Steveninck, 2000; Schreiber, 2000; Balasubramanian, Kimber, & Berry, 2001; de Polavieja, in press).

We set out to investigate the relationship between energy and information at the level of the discrete molecular events that generate cell signals. Ultimately, information is transmitted by the activation and deactivation of signaling molecules. These are generally single proteins or small complexes that respond to changes in electrical, chemical, or mechanical potential. Familiar neural examples are the opening of an ion channel, the binding of a ligand to a receptor, the activation of a G-protein, and vesicle exocytosis. These events involve the expenditure of energy—for example, to restore ions that flow across the membrane, restore G-proteins to the inactive (G-GDP) state, and remove and recycle neurotransmitter. The ability of these events to transmit information is limited by their stochasticity (Laughlin, 1989; White, Rubinstein, & Kay, 2000). This uncertainty reduces reliability and hence the quantity of transmitted information. To increase information, one must increase the number of events used to transmit the signal; this, in turn, increases the consumption of energy. We investigate this fundamental relationship between information and energy in molecular signaling systems by developing a simple model within the context of neural information processing: a population of ion channels that responds to an input by changing their open probability. We derive the quantity of information transmitted by the population of channels and demonstrate how information varies as a function of the properties of the input and the number of channels in the population. We identify optima that maximize the ratio between transmitted information and cost. These optima depend on the input statistics and, as with spike codes (Levy & Baxter, 1996), the ratio between the costs of generating signals (the signaling cost) and the cost of constructing the system, maintaining it in a state of readiness and providing it with an input (the fixed costs).

The article is organized as follows. Section 2 introduces the model for a system that transmits information with discrete stochastic signaling events.

In section 3, we define measures of information transfer, energy consumption, and metabolic efficiency for such a system. In section 4, we analyze the dependence of energy efficiency on the number of stochastic units for gaussian input distributions. The influence of additional noise sources is studied in section 5. In section 6, we look at the energy efficiency of combinations of systems, and in section 7, we derive optimal input distributions and show how energy efficiency depends on the number of stochastic units when the distribution of inputs is optimal. Finally, in section 8 we conclude our investigation with an extensive discussion.

2 The Model

We consider an information transmission system with input x and output k. The system has N identical, independent units that are activated and deactivated stochastically. The input x directly controls the probability that a unit is activated. The number of activated units, k, constitutes the output of the system.

Because realistic physical inputs are bounded in magnitude, any given distribution of inputs can be mapped in a one-to-one fashion onto the activation probabilities of units, within the interval $[0; 1]$. We therefore assume, without loss of generality, that $x \in [0; 1]$ is equivalent to the probability of being in an activated state. Consequently, the conditional probability that a given input x activates k units is given by a binomial distribution,

$$p(k|x) = \binom{N}{k} x^k (1 - x)^{N-k}. \tag{2.1}$$

The variance $\sigma^2_{k|x}$ of this binomial probability distribution,

$$\sigma^2_{k|x} = Nx(1 - x), \tag{2.2}$$

is a measure for the system's transduction accuracy, defining the magnitude of the noise. Note that $\sigma^2_{k|x}$ depends on both the number of available units N and the input x.

In an equivalent interpretation, the model can also be considered as a linear input-output system,

$$k = Nx + \eta(N, x), \tag{2.3}$$

where Nx is the "deterministic" component of the output and $\eta(N, x)$ represents the noise due to the stochasticity of the units. The noise distribution corresponds to $p(k|x)$ shifted to have zero mean; its variance is therefore given by $\sigma^2_{k|x}$. Thus, we see that the input x specifies a noise-free output $N \cdot x$, to which the noise η is added, yielding the integer-valued output k.
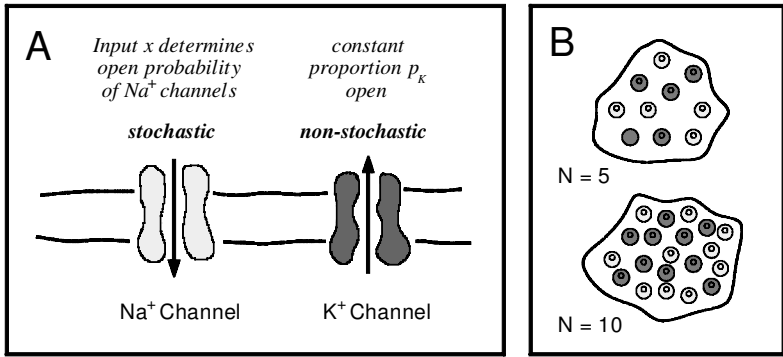
Figure 1: (A) Schematic view of the membrane model. The input x directly determines the probability that sodium channels are open. In contrast to the stochastic sodium channels, potassium channels are considered nonstochastic at this stage of analysis; independent of the input, a constant fraction p_K is open. (B) Schematic view of two model signaling systems with $N = 5$ and $N = 10$ sodium channels, respectively. Note that the ratio of sodium to potassium channels is kept constant (here $N_K/N = 1$).

2.1 Implementing the Model. For concreteness, we have chosen to implement the model in the context of a basic neural signaling mechanism. A membrane contains two populations of voltage-insensitive channels: one controlling inward current and the other outward current (see Figure 1). For convenience, we refer to these as sodium and potassium channels, respectively. There are N sodium channels and N_K potassium channels.

In our analysis, an input, x, produces a voltage response by changing the open probability of the set of sodium channels, which take the role of the stochastic signaling units. The input x could be derived from a variety of sources, both external, such as light, and internal, such as synaptically released transmitter. But regardless of its origins, the input is assumed to be unambiguously represented by the open probability of sodium channels. For simplicity, the second set of ion channels—the potassium channels—is considered to be input independent and noise free. Thus, a fixed proportion of potassium channels is kept open, regardless of the size of the input.

The input signal specifies the voltage output signal by directly determining the probability that sodium channels are open. Thus, a given input value x, presented in a small time interval $\Delta\tau$, will result in k open sodium channels. Note that because of channel noise, the number of open sodium channels k will vary with different presentations of the same input. The state of the model system, on the other hand, is given by the number of open sodium channels k and translates uniquely into a voltage output if we neglect the influence of membrane capacitance. Conversely, if we know the

output voltage V, we can directly infer the number of open sodium channels k. Both variables are therefore equivalent from an information-theoretic point of view. By working in the channel domain (i.e., by taking the number of open sodium channels k as a measure of the system's output), we can simplify our analysis and avoid the nonlinear relationship of conductance, current, and voltage.

Note that to achieve the linear relationship between the input x and the (average) output k, the channels are not voltage sensitive. For simplicity, we do not study the effects of membrane capacitance on the time course of the voltage response, assuming that the signal variation is slow in comparison to the timescale of the effects introduced by membrane capacitance. Nor do we analyze the effects of channel kinetics. By working at a fundamental level, the mapping of the input onto discrete molecular events, we can investigate a simple model of general validity.

3 Calculating the Metabolic Efficiency

We define the metabolic efficiency of a signaling system as the ratio between the amount of transmitted information I and the metabolic energy E required. Efficiency I/E thus can be expressed in bits per energy unit. (For other efficiency measures, see section A.3.) Both information transfer I and metabolic energy (or cost) E depend on the number of available signaling units—the number of channels, N. To investigate the relationship between efficiency and the number of stochastic units, we drive the model system with a fixed distribution of inputs, $p_x(x)$, and vary N, the size of the population of sodium channels used for signaling. This is equivalent to changing the channel density of our model membrane while maintaining the same input. To ensure that systems with different values of N, the number of sodium channels, produce the same mean voltage output in response to a given input, x, the population of potassium channels, N_K, is set to a constant proportion of N. Under these conditions, we can now calculate how the information transmitted, I, and the energy consumed, E, vary for different numbers of channels.

3.1 Information Transfer. Consider the transmission of a single signal. The model system receives an input, x, drawn from the input distribution, $p_x(x)$, and produces a single output, k. According to Shannon, the information transmitted by the system is given by

$$I[N; p_x(\cdot)] = \sum_{k=0}^{N} \int_0^1 dx\, p(k|x)p_x(x) \log_2 \left[\frac{p(k|x)}{p_k(k)} \right], \tag{3.1}$$

and depends on the input distribution $p_x(x)$ and, via the binomial distribution of channel activation $p(k|x)$ (see equation 2.1), also on the number

of available units, N. For a continuously signaling neuron, this is equivalent to the response during a discrete time bin of duration $\Delta\tau$, and I is the rate of transmission in bits/$\Delta\tau$. In this article, we study two different scenarios—gaussian input distributions and input distributions that maximize the information transfer—both in the presence of noise caused by stochastic units (e.g., channel noise).

3.2 Energy. The total amount of energy required for the maintenance of a signaling system and the transmission of a signal is given by

$$E[N; p_x(\cdot)] = b + \int_0^1 p_x(x)e(x, N)\,dx, \tag{3.2}$$

where b is the fixed cost, and $e(x, N)$ characterizes the required energy as a function of the input x and the number of stochastic signaling units N. Thus, we classify the metabolic costs into two groups: costs that are independent of N and costs that depend on the total number of channels, N. For simplicity, we assume that the latter costs are dominated by the energy used to generate signals (in this case, restoring the current that flows through ion channels), and we neglect the energy required to maintain channels. The first group of costs, b, relates to costs that have to be met in the absence of signals, such as the synthesis of proteins and lipids. These costs are therefore called fixed costs and are constant with respect to x and N. Because we have set up our systems to produce identical mean values of the voltage output signal given x (by fixing the ratio N/N_K), the function $e(N, x)$ is separable into the variables N and x (see section A.1),

$$e(x, N) = N\tilde{e}(x), \tag{3.3}$$

so that the signaling-related total energy consumption rises linearly with N. The function $\tilde{e}(x)$ increases monotonically between $\tilde{e}(0) = 0$ and $\tilde{e}(1) = e_{ch}$, where e_{ch} denotes the energy cost associated with a single open sodium channel. (The precise form of $\tilde{e}(x)$ is derived in section A.1.) Rescaling the measurement units of energy, we will from now on set $e_{ch} = 1$. Altogether, the total energy reads

$$E[N; p_x(\cdot)] = b + \overline{\tilde{e}(x)}N, \tag{3.4}$$

where $\overline{\tilde{e}(x)}$ is the average signal-related energy requirement of one stochastic unit and the average is taken with respect to the input distribution $p_x(x)$. In the first part of the analysis, where we analyze energy-efficient numbers of channels, we make the simplifying assumption that the average cost per channel, $\overline{\tilde{e}(x)}$, is approximately equal to the mean of the input, $\overline{x}$. Note that the energy E is defined as a measure of cost for one time unit $\Delta\tau$, just as I is the measure of information transfer in $\Delta\tau$.

4 Gaussian Input Distributions

4.1 Information Transfer. We focus on gaussian inputs first, because according to the central limit theorem, they are a reasonable description of signals composed of many independent subsignals, and they also allow an analytic expression of information transfer. To confine the bulk of the gaussian distributions within the input interval [0; 1], the mean, $\overline{x}$, and the variance, σ_x^2, are chosen such that the distance from the mean to the interval borders is always larger than $3\sigma_x$. Values falling outside the interval [0; 1] are cut off, and the distribution is then normalized to unity. Numerical simulations (see section A.2) show that the effects of this procedure on the results are negligible.

The information transfer I (Shannon & Weaver, 1949) per unit time for a linear system with additive gaussian noise and gaussian inputs is given by

$$I = \frac{1}{2} \log_2(1 + SNR), \tag{4.1}$$

where SNR denotes the signal-to-noise ratio. It is defined as the ratio between the signal variance, σ_x^2, and the effective noise variance, $\sigma_{k|x}^2/N^2$. If two criteria are met—first, the binomial noise $\eta(N, x)$ can be approximated by a gaussian and, second, within the regime of most likely inputs x, changes in the noise variance $\sigma_{x|k}^2$ are negligible—the following equation gives a reasonably good approximation of the information transfer of our model system:

$$I = \frac{1}{2} \log_2 \left(1 + \frac{N\sigma_x^2}{\overline{x}(1 - \overline{x})} \right). \tag{4.2}$$

This is the case for large N and a restriction to the gaussian inputs described above. Numerical tests (see section A.2) show that the deviation between the real information transfer with N stochastic signaling units and the information transfer given by equation 4.2 is very small.

4.2 Efficiency. For the efficiency, defined as I/E, we obtain the following expression:

$$\frac{I}{E} = \frac{1}{2\,(\overline{x}N + b)} \log_2 \left(1 + \frac{N\sigma_x^2}{\overline{x}(1 - \overline{x})} \right). \tag{4.3}$$

In Figure 2, efficiency curves I/E are depicted as a function of N for three different values of fixed costs b. The energy efficiency exhibits an optimum for all curves. Signal transmission with numbers of channels N within the range of the optimum is especially energy efficient. The position of the optimum depends strongly on the size of the fixed cost b relative to the average cost of a single channel $\overline{e(x)} \approx \overline{x}$. Figure 2 displays the dependence of the
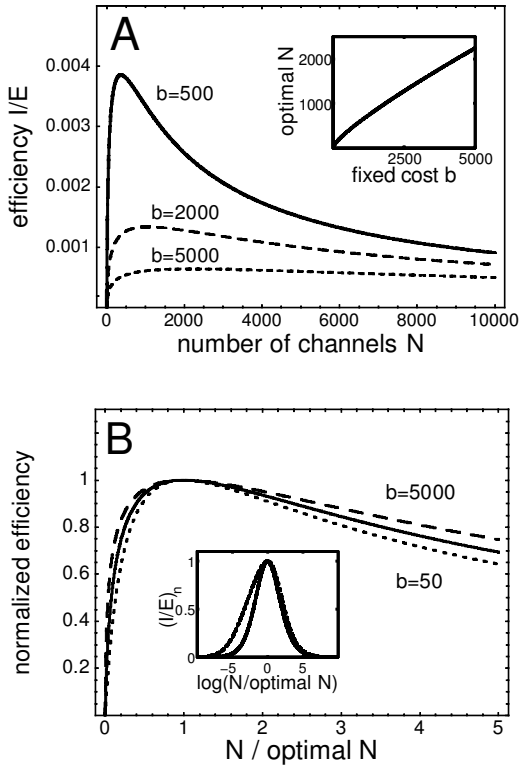
Figure 2: (A) Efficiency I/E as a function of N for three different fixed costs. From top to bottom, the fixed cost b is equivalent to the cost of 500, 2000, and 5000 open channels. The optimal N shifts to larger N with rising fixed cost b (see also the inset), and the efficiency curves become wider. (B) Efficiency curves for fixed costs of $b = 50$, $b = 500$, and $b = 5000$ rescaled, so that the maximum corresponds to (1,1). The inset shows the same data on a logarithmic scale. These rescaled curves, $(I/E)_n$, are similar in shape, independent of the size of b. For all graphs, the parameters of the input distribution $p_x(x)$ are $\sigma_x = 0.16$ and $\bar{x} = 0.5$.

optimal number of channels N on the fixed cost b. The most efficient number of channels increases approximately linearly with the size of the fixed cost, although close inspection reveals that the slope is steeper for smaller fixed costs and shallower in the range of higher b. If $b = 0$, the most efficient system capable of transmitting information uses one channel. The average costs of the most energy-efficient population of channels, employing N_{opt} channels, are given by $N_{opt}\bar{x}$. Therefore, the ratio between these signaling costs at $\bar{x} = 0.5$, and the fixed cost b is approximately 1:4 in the example depicted. This ratio however, which can be derived from the slope of the

b-N_{opt} curve in the inset of Figure 2A, strongly depends on the input distribution. The analysis for gaussian input distributions shows that the ratio of N_{opt} to b increases with decreasing input variance, σ_x^2 (results not shown).

Remarkably, the efficiency curves rise very steeply for small N and, after passing the optimum, decrease only gradually. This characteristic does not strongly depend on the size of b, as shown in Figure 2B. It is thus very uneconomical to use a number of channels sufficiently far below the optimum, whereas above, there is a broad range of high efficiency. In this range, a cell could adjust its number of channels to the amount of information needed (e.g., add more channels), without losing much efficiency. However, as the inset to Figure 2B indicates, increasing the number of channels by a given factor has a similar effect on efficiency as decreasing them by the same factor.

5 Additional Noise Sources

The most energy-efficient number of channels is influenced by the size of additional noise. This might be noise added to the input (additive input noise) or internal noise generated within the signaling system independent of the activation of sodium channels (input-independent internal noise).

5.1 Additive Input Noise.
If the input contains gaussian noise of a fixed variance $\langle \eta_x^2 \rangle$ that is not correlated with the signal, the variances of the signal and the noise add, yielding the modified SNR at the output

$$SNR = \frac{N\sigma_x^2}{N\langle \eta_x^2 \rangle + \overline{x}(1 - \overline{x})}. \tag{5.1}$$

The additional noise η_x decreases the SNR and, consequently, the information transfer. For $N \rightarrow \infty$, the SNR converges to $\sigma_x^2/\langle \eta_x^2 \rangle$, and thus sets an upper limit to the information transfer. Figure 3 shows that it is more efficient to operate at lower numbers of channels in the presence of additional signal noise.

5.2 Input-Independent Internal Noise.
To exemplify internal noise, we consider an additional population of sodium channels that is not influenced by the input x but rather has a fixed open probability p', though it contributes to the noise. Assuming a fixed ratio N/N' between the total number of the original input-dependent sodium channels N and the total number of these new input-independent sodium channels N', the voltage output is determined by the sum of open channels from both populations $k + k'$. Because noise from both populations is uncorrelated, the SNR reads

$$SNR = \frac{N^2\sigma_x^2}{N^2\langle \eta_x^2 \rangle + N\overline{x}(1 - \overline{x}) + N'p'(1 - p')}, \tag{5.2}$$
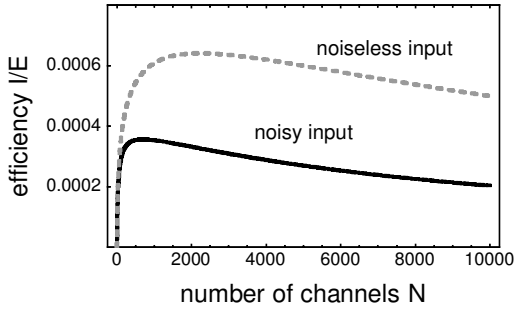
Figure 3: Efficiency I/E as function of the number of signaling units, N, for the case of a noisy input signal (solid line). For comparison, we also replot the case where there is no input noise (dotted line), as in Figure 2A. For both curves, the variance of the input distribution equals $\sigma_x^2 = 0.16^2$, $\overline{x} = 0.5$, and $b = 5000$. The noise variance of the signal is $\langle \eta_x^2 \rangle = 0.04^2$.

where $N'p'(1 - p')$ represents the noise variance of the input-independent population of channels. The efficiency of the signaling system is thus further decreased for all N, and the efficiency optimum, N_{opt}, is shifted to lower values.

Other noise sources, such as leak channels and additional synaptic inputs, will also lower the SNR. Therefore, they will reduce efficiency and influence the optimum number of channels.

6 Efficiency in Systems Combining Several Mechanisms

Signaling mechanisms do not act in isolation; they are usually organized into systems in which one mechanism drives another, either within a cell or between cells. The relationship between information transfer and cost of each mechanism determines optimal patterns of investment in signaling units across the whole system, as we will demonstrate with some simple examples.

First, consider two signaling mechanisms in series (see Figure 4). Cell 1 uses N_1 channels to convert the input x into the output k_1, which, in turn, drives cell 2, which uses N_2 channels, to produce the output k_2. From equation 2.3, we know that $k_1 = N_1 x + \eta_1$, where Nx is the signal and η_1 is the noise generated by the random activity of channels. Because we define an input in terms of a probability distribution of signals, ranging from 0 to 1, the output k_1 of cell 1 should be normalized by N_1, so that the input to cell 2 is k_1/N_1. Note that, for simplicity, we are neglecting nonlinearities in signal transfer within a cell, as, for example, in neurotransmitter release. As a consequence, the mean open probability in both cells is the same, but its variance differs. The output of cell 2 is given by $k_2 = N_2 x_2 + \eta_2$, where η_2
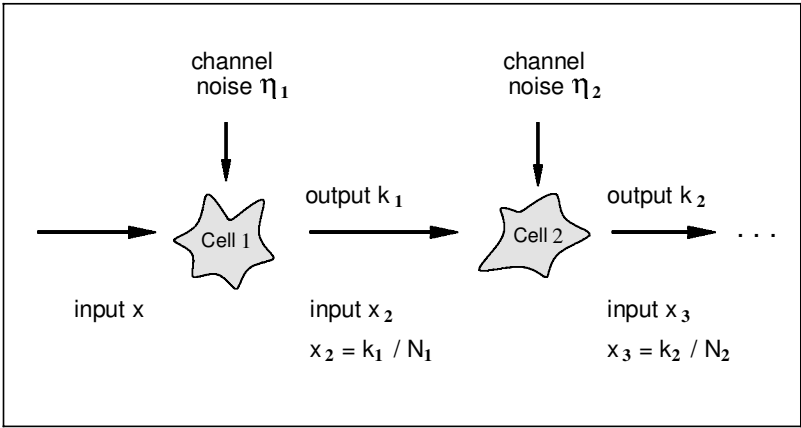
Figure 4: Schematic view of the cell arrangement. The normalized output of cell 1 serves as input for cell 2. Both cells are subject to channel noise η_1 and η_2, respectively.

is the additive channel noise of cell 2. Therefore, the information transfer from an input signal x, with mean $\overline{x}$ and variance σ_x^2, to the output k_2 can be approximated by Shannon's formula as

$$I = \frac{1}{2} \log_2 \left(1 + \frac{N_1 N_2 \sigma_x^2}{(N_1 + N_2)\overline{x}(1 - \overline{x})} \right). \tag{6.1}$$

The cost of information transfer through the two-cell system is

$$E = e_{ch1}\overline{x}N_1 + e_{ch2}\overline{x}N_2 + b_1 + b_2, \tag{6.2}$$

where b_1 and b_2 are the fixed metabolic costs of the cells and $e_{ch1/2}\overline{x}$ the costs per open channel.

If we introduce an effective number of channels, $N_{eff}^i = N_1 N_2/(N_1 + N_2)$, for the information transfer and $N_{eff}^E = N_1 + N_2$ for the metabolic cost, the equations for I and E correspond to those of the single-cell case—equations 4.2, and 3.4, respectively. For simplicity, the cost per open channel is set to unity for both cells. Because $N_{eff}^E \geq N_{eff}^i$ for all nonnegative N_1 and N_2, the information transfer increases more slowly with the number of channels than the cost, cutting down efficiency. Thus, a two-cell system requires more channels to transmit the same amount of information and is therefore less efficient than a single cell, even if the fixed cost of a cell in the two-cell model is only half the cost of the single cell. Consequently, in a metabolically efficient nervous system, serial connections should be made only when signal processing is required.

Furthermore, in an energy-efficient system, the cost of one mechanism influences the amount of energy that should be consumed by another. Such considerations are important when one set of signaling events is more costly than another. This can be demonstrated by incorporating the cost of an upstream mechanism into the fixed cost of a downstream mechanism (i.e., the cost of a mechanism includes the cost of providing it with a signal). Here we can define E as

$$E = e_{ch2}\overline{x}N_2 + b_2^* \qquad \text{with} \qquad b_2^* = e_{ch1}\overline{x}N_1 + b_1 + b_2. \tag{6.3}$$

As we have seen with one mechanism, an increase in fixed cost raises the optimal number of channels. Therefore, when the cost of an upstream mechanism is high, one improves energy efficiency by using more units downstream. The precise pattern of investment required for optimal performance will depend on the relative costs (fixed and signaling) of every mechanism (e.g., channels) and the characteristics of the input signal.

7 Limits to the Achievable Efficiency

Information transfer and efficiency depend on the distribution of input signals. In the previous sections, we have considered gaussian inputs. We now calculate the input distribution that maximizes information transfer I given a limited energy budget E and a particular number of channels N. The efficiency I/E reached gives an upper bound on the efficiency the system can achieve for given E and N. Although the nervous system has less influence on the distribution of external signals, it is able to shape its internal signals to optimize information transfer (Laughlin, 1981).

The optimal input distribution, $p_x^{opt}(x)$, and the maximum information transfer (the information capacity C_N) of a system with N stochastic units can be obtained by the Blahut-Arimoto algorithm (Arimoto, 1972; Blahut, 1972), which is described in further detail in section A.4. Given the noise distribution $p(k|x)$, the algorithm yields a numerical solution to the optimization of the input distribution $p_x(x)$, maximizing the information transfer and minimizing the metabolic cost. This algorithm has been applied by Balasubramanian et al. (2001) and de Polavieja (in press) to study the metabolic efficiency of spike coding.

7.1 Optimal Input Distributions. For a given number of channels N, a given fixed cost b, and a given cost function depending on the input $\tilde{e}(x)$, the energy $E = N\overline{\tilde{e}(x)} + b$ used by the system depends exclusively on the input distribution $p_x(x)$. If the available energy budget is sufficiently large, energy constraints do not influence the shape of the optimal input distribution, and the information capacity C_N of the system reaches its maximum (see Figure 5E, point A). The optimum input distribution turns out to be symmetrical, with inputs from the midregion around $x = 0.5$ drawn less of-
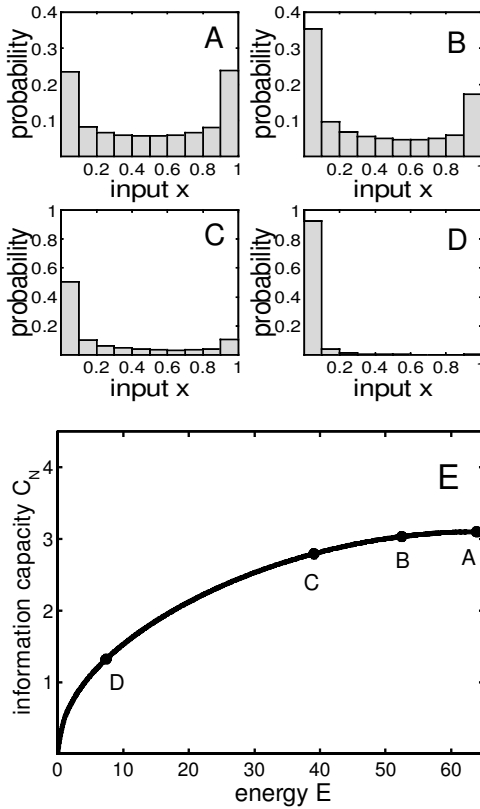
Figure 5: (A)–(D) Optimal input distributions for different energy budgets E with $b = 0$ and $N = 100$. The distributions were calculated numerically and are discretized to a resolution $\Delta x = 0.01$. All distributions show small values for inputs around $x = 0.5$ where the noise variance $\sigma^2_{k|x}$ is highest. With decreasing energy budget, the distributions become less symmetrical, preferring low inputs. (E) Information capacity $C_{N=100}$ depending on the energy budget (for $b = 0$). The points mark the location of the input distributions shown in A–D in the energy capacity space.

ten, whereas inputs at the boundaries 0 and 1 are preferred (see Figure 5A). This result is very intuitive, when we take the dependence of the noise variance, $\sigma^2_{k|x}$, as defined in equation 2.2, on the input x into account. The noise variance is symmetrical as well, showing a maximum at $x = 0.5$ and falling off toward $x = 0$ and $x = 1$. Thus, an input distribution that is optimal from the point of view of information transfer, leaving metabolic considerations aside for a moment, favors less noisy input values over noisier ones.
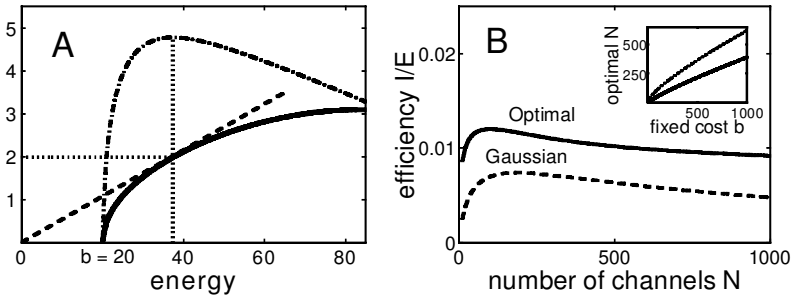
Figure 6: (A) Capacity $C_{N=100}$ (of Figure 5E) as a function of energy E for a fixed cost $b = 20$ (solid line). The capacity curve is simply shifted along the energy axis by the value of b. The efficiency C_N/E as a function of E is shown as a dash-dotted line. The maximum efficiency $(C_N/E)_{\max}$ is given at the point of the $C_N(E)$ curve whose tangent (dashed line) intersects with the origin. (B) Achievable efficiency $(C_N/E)_{\max}$ as a function of N for a fixed cost $b = 200$ (solid line). For comparison, the efficiency I/E for a gaussian input distribution ($\bar{x} = 0.5$, $\sigma_x = 0.16$) is also shown (dashed line). The inset depicts the optimal number of channels N_{opt} as a function of b.

Limiting the energy that can be used by a system of N units, however, changes the optimal input distribution, $p_x(x)$, and destroys the symmetry. As we reduce the energy budget, values neighboring $x = 0$ are increasingly preferred and the costly values approaching $x = 1$ are avoided (see Figures 5B–5D). This asymmetry reduces the information capacity C_N. Thus, efficient use of a restricted budget requires a cell to keep most of its units deactivated. For our simple model, this is equivalent to maintaining the membrane close to resting potential by keeping most of its sodium channels closed.

The fixed cost, b, is a metabolic cost independent of the value of the input x and cannot be avoided. Consequently, it does not influence the shape of the energy-capacity curve of a system. Adding the fixed cost b results merely in a horizontal translation of the energy-capacity curve (see Figure 6A). Here as well, the shape of the input distribution changes with the value of E.

7.2 Efficiency. Having obtained the dependence of the information capacity on the energy used for a given N, we can also derive energy efficiency C_N/E as a function of the energy E. Of particular interest is the maximum value $(C_N/E)_{\max} = \max_E\{C_N/E\}$, giving the optimal efficiency for a fixed value of N, which is achieved by a specific input distribution $p_x(x)$. Note that this efficiency gives the upper bound to the achievable efficiency in our system and therefore cannot be surpassed by any other input distribution $p_x(x)$.

At the maximum of C_N/E, the first derivative with respect to E is zero,

$$\frac{\partial}{\partial E}\left(\frac{C_N}{E}\right) = 0, \tag{7.1}$$

which can be transformed to give

$$\frac{\partial C_N}{\partial E} \cdot E = C_N. \tag{7.2}$$

Thus, geometrically, the maximum value $(C_N/E)_{max}$ corresponds to the point on the capacity graph whose tangent intersects the origin (see Figure 6A). The slope of the tangent is given by $(C_N/E)_{max}$ itself, so that the optimal efficiency $(C_N/E)_{max}$ decreases with increasing b, as can be inferred from Figure 6A by shifting the $C_N(E)$ curve to the right.

Figure 6B shows the optimal efficiency $(C_N/E)_{max}$ as a function of the number of channels N for a fixed cost $b = 200$. For comparison, we also show the efficiency I/E obtained for a gaussian input distribution. The optimal input distribution surpasses the gaussian distribution roughly by a factor of two in this case.

In conclusion, as with gaussian input distributions, we can derive the most efficient number of channels N_{opt} as a function of the fixed cost. The use of optimal input distributions reduces N_{opt} but, as with gaussian inputs, N_{opt} rises approximately linearly with the basic cost, b, as illustrated in the inset of Figure 6B.

8 Discussion

A growing number of studies of neural coding suggest that the consumption of metabolic energy constrains neural function (Laughlin, 2001). Comparative studies indicate that the mammalian brain is an expensive tissue whose evolution, development, and function have been shaped by the availability of metabolic energy (Aiello & Wheeler, 1995; Martin, 1996). The human brain accounts for about 20% of an adult's resting metabolic rate. In children, this proportion can reach 50% and in electric fish 60% (Rolfe & Brown, 1997; Nilsson, 1996). Because much of this energy is used to generate and transmit signals (Siesjö, 1978; Ames, 2000), these levels of consumption have the potential to constrain neural computation by placing an upper bound on synaptic drive and spike rate (Attwell & Laughlin, 2001).

Current evidence suggests that the advantages of an energy-efficient nervous system are not confined to larger animals with highly evolved brains. In general, the specific metabolic rate of brains is 10 to 30 times the average for the whole animal, at rest (Lutz & Nilsson, 1994). In addition, for both vertebrates and insects, the metabolic demands of the brain are more acute in smaller animals because the ratio of brain mass to total body mass de-

creases as body mass increases (Martin, 1996; Kern, 1985). Moreover, insect species with similar body masses exhibit large differences in the total mass of the brain and in the masses or relative volumes of different brain areas, and these differences correlate with behavior and ecology (Kern, 1985; Gronenberg & Liebig, 1999; Laughlin et al., 2000). Significant changes have been observed among individuals of a single species. In the ant *Harpegnathos*, the workers are usually visually guided foragers. However, when young workers are inseminated and begin to lay eggs following the death of their queen, their optic lobes are reduced by 20% (Gronenberg & Liebig, 1999). These observations suggest that the reduction of neural energy consumption is also a significant factor in the evolution of small brains.

The relationship between energy consumption and the ability of neurons to transmit information suggests that nervous systems have evolved a number of ways of increasing energy efficiency. These methods include redundancy reduction (Laughlin et al., 1998), the mix of analog and digital operations found in cortical neurons (Sarpeshkar, 1998), appropriate distributions of interspike intervals (Baddeley et al., 1997; Balasubramanian et al., 2001; de Polavieja, 2001), and distributed spike codes (Levy & Baxter, 1996). We have extended these previous theoretical investigations to a level that is more elementary than the analysis of signaling with spikes: the representation of information by populations of ion channels. Thus, as independently advocated by Abshire and Andreou (2001), we have analyzed the energy efficiency of information transmission at the level of its implementation by molecular mechanisms.

We have estimated the amount of information transmitted by a population of voltage-insensitive channels when their open probability is determined by a defined input. These channels typify the general case of signaling with stochastic events. Consequently, our analysis is also applicable to many other forms of molecular signaling (e.g., the binding of a ligand to a receptor) and to synaptic transmission (e.g., the release of a synaptic vesicle according to Poisson or binomial statistics). Our theoretical results verify a well-known trend: the amount of information carried by a population of channels increases with the size of the population because random fluctuations are averaged out, as observed in blowfly photoreceptors and their synapses (de Ruyter van Steveninck, Lewen, Strong, Koberle, & Bialek, 1997; Laughlin et al., 1998; Abshire & Andreou, 2000) and demonstrated by models of synaptic transmission to cortical neurons (Zador, 1998). However, increasing the number of channels in the population increases both the level of redundancy and the energy used for transmission, leading to changes in metabolic efficiency (Laughlin et al., 1998).

Following Levy and Baxter (1996), we have chosen to discuss efficiency as the ratio between the number of bits of information transmitted and the energy consumed. However, our analysis also provides a mathematical framework to describe energy efficiency from the more general point of view of maximizing information transfer and minimizing the metabolic cost

(maximizing $I - sE$, where s describes the importance of energy minimization over information maximization), as briefly outlined in section A.3. We distinguish two energy costs: the cost of generating signals and the fixed cost of keeping a signaling system in a state of readiness. The signaling cost is derived from the current flow through ion channels. Under the assumptions of the model, the signaling cost increases with the number of channels. This simple linear relationship can be easily applied to other forms of signaling, such as protein phosphorylation or the turnover of transmitter and second messenger. In the absence of data on the costs of constructing and maintaining a population of channels in a membrane, we again follow Levy and Baxter (1996). The fixed cost is expressed in the same units as the signaling cost and is varied to establish its effect on efficiency.

The analysis demonstrates that energy efficiency is maximized by using a specific number of channels. These optima depend on a number of important factors: the fixed cost of the system, the cost of signaling, the reliability of the input, the amount of noise generated by other intrinsic mechanisms, the cost of upstream and downstream signaling mechanisms, and the distribution of the input signals provided by upstream mechanisms. Each of these factors is involved in neural processing.

The fixed cost of building and maintaining the cell in a physiological state within which the ion channels operate is a dominant factor. When the fixed cost increases, the optimum system increases the energy invested in signaling by increasing the number of channels (see Figure 2A). Levy and Baxter (1996) discovered the same effect in their analysis of energy-efficient spike trains. This may well be a general property of energy-efficient systems because when a signaling system is expensive to make and maintain (i.e., the ratio between fixed and signaling costs is high), it pays to increase the return on the fixed investment by transmitting more bits. This takes more channels and more signaling energy. For the example shown, which operates with a broad input distribution and a mean open probability of 50%, the optimum population of channels has a peak energy consumption (all channels open) that is approximately half the fixed cost (see Figure 2A). The relationship between the energy spent on signaling by the channels and the fixed cost varies with the distribution of inputs. For input distributions that make a reasonably broad use of possible open probabilities, the ratio between signaling costs and fixed costs lies approximately in the range between 1:4 and 1:1. It is difficult to judge whether populations of neuronal ion channels follow this pattern because data about the ratio of fixed costs to signaling costs in single cells are not available. However, in more complicated systems, the proportion of energy devoted to signaling is in the predicted range. For the whole mammalian brain, signaling has been linked to approximately 50% of the metabolic rate (Siesjö, 1978), and in cortical gray matter this rises to 75% (Attwell & Laughlin, 2001).

We are aware that there are a number of additional factors, not accounted for by our model, that will influence the ratio of signaling costs to fixed costs

in nervous systems. For example, our analysis underestimates the total energy usage by the brain because it is confined to a single operation: the generation of a voltage signal by channels. Within neural systems, signals must be transmitted over considerable distances, and various computations must be performed. These extra operations take extra energy. Along these lines, the transmission of signals along axons, in the form of action potentials, accounts for over 40% of the signaling cost of cortical gray matter (Attwell & Laughlin, 2001).

Noise at the input reduces the optimum number of channels (see Figure 3) because it reduces the effect of channel fluctuations on the total noise power. There is some evidence that nervous systems reduce the number of signaling events in response to a less reliable input. In the blowfly retina, the SNR of chromatic signals is lower than that of achromatic signals, and a class of interneurons involved in the chromatic pathway uses fewer synapses than comparable achromatic interneurons (Anderson & Laughlin, 2000). Considering a chain of signaling mechanisms allows us to study networks where the output from one population of channels defines the SNR at the input of the next. As a result, when analog signals are transferred from one mechanism to another, the noise accumulates stage by stage (Sarpeshkar, 1998). The chain describes how this buildup of noise reduces metabolic efficiency. Given this reduction, an energy-efficient system should connect one population of channels (or synapses) to another in a serial manner only when information is actually processed, not when it is merely transmitted. Where signals must be repeatedly amplified to avoid attenuation, pulses should be used to resist the buildup of analog noise (Sarpeshkar, 1998). These design strategies are hypothetical. The energy savings that are made by restricting the number of serial or convergent analog processes and converting analog signals to spikes (Sarpeshkar, 1998; Laughlin et al., 2000) have yet to be demonstrated in a neural circuit.

Our analysis suggests that when transmission and processing involve several types of signaling events (e.g., calcium channels, synaptic vesicles, and ligand gated channels at a synapse), it is advantageous to use more of the less expensive events and fewer of the more expensive. This distribution is analogous to the pattern of investment in bones of a mammal's leg. Proximal bones are thicker than distal bones because they are less costly per unit mass (they move less during a stride). The thickening of distal bones is adjusted to optimize the ratio between the probability of fracture and cost for the limb as a whole (Alexander, 1997).

Finally, the probability distribution of the input signal has a large effect on efficiency. On both an evolutionary timescale as well as on the timescale of physiological adaptation processes, the way an external signal with specific statistical properties is transmitted could therefore be optimized by mapping the external signal distribution on an efficient distribution of probabilities of channels to be open (which we call the input distribution). More

importantly, the internal signals passed on from one mechanism to the next could be shaped such that the signal distributions employed will enhance the efficiency of information transfer. The Blahut-Arimoto algorithm yields input distributions that optimize the amount of information transferred under a cost constraint and has been successfully applied to spike codes (Balasubramanian et al., 2001; de Polavieja, 2001). Our application shows how inputs can be mapped onto the probabilities of activating signaling molecules to maximize the metabolic efficiency of analog signaling. The improvement over gaussian inputs is greater than 50% and is achieved by two means. First, signaling avoids the midregion of activation probabilities where, according to binomial statistics, the noise variance is high. Second, signaling avoids the expensive situation of having a high probability of opening channels in favor of the energetically cheaper low-probability condition, similar to the results that a metabolically efficient spike code avoids high rates (Baddeley et al., 1997; Balasubramanian et al., 2001). Our analysis suggests that an efficient population of sodium and potassium channels usually operates close to resting potential, with most of its sodium channels closed, but infrequently switches to opening most of its sodium channels. In other words, there is a tendency toward using a combination of numerous small signals close to a low resting potential and less frequent voltage peaks.

In conclusion, we have analyzed the energy efficiency of a simple biological system that represents analog signals with stochastic signaling events, such as ion channel activation. Optimum configurations depend on the basic physics that connects information to energy (the dependency of noise, redundancy, and cost on the number of signaling events) and basic economics (the role played by fixed costs in determining optimum levels of expenditure on throughput). Given this fundamental basis, the principles that we have demonstrated are likely to apply to other systems. In particular, we have shown that energy efficiency is a property of both the component mechanism and the system in which it operates. To assess a single population of ion channels, we had to consider the fixed cost, the distribution of signal and noise in the input, and additional noise sources. After connecting two populations of ion channels in series, we had to add the relative costs of the two mechanisms and the noise fed from one population to the next to our list of factors. Energy efficiency is achieved by matching the mechanism to its signal or, for optimum input distributions, the signal to the mechanism. Matching is a characteristic of designs that maximize the quantity of information coded by single cells, regardless of cost. To achieve this form of efficiency, neurons exploit the adaptability of their cellular and molecular mechanisms (Laughlin, 1994). The extent to which the numbers of channels and synapses used by neurons, and their transfer functions, are regulated for metabolic efficiency remains to be seen. The analysis presented here provides a starting point that can guide further experimental and theoretical work.

Appendix

A.1 Energetic Cost of Inputs. For the channel model, the average energetic cost per unit time $e(x, N)$ is a function of the input x and the number of sodium channels N. The sodium and potassium currents, i_{Na} and i_K, depend on the reversal potentials E_{Na} and E_K, the membrane potential V, as well as on the conductances of the sodium and potassium channels, respectively,

$$i_{Na}(x) = N g_{Na0} x (V(x) - E_{Na}), \tag{A.1}$$

$$i_K(x) = N_K g_{K0} p_K (V(x) - E_K). \tag{A.2}$$

The vigorous electrogenic pump extrudes three sodium ions and takes up two potassium ions for every ATP molecule hydrolyzed. The pump current i_{pump} equals $i_K/2$, assuming that the pump maintains the internal potassium concentration by accumulating potassium ions at a rate equal to the outward potassium current, i_K. Equating all currents across the membrane gives

$$i_{Na}(x) + i_K(x) + i_{pump}(x) = i_{Na}(x) + \frac{3}{2} i_K(x) = 0. \tag{A.3}$$

The energetic cost $e(x, N)$ is proportional to the pump current, i_{pump}, so that we define

$$e(x, N) = c \cdot i_{pump}(x) \tag{A.4}$$

$$= cN \cdot \frac{g_{K0} p_K \left(\frac{N_K}{N} \right) g_{Na0} (E_{Na} - E_K) x}{3 g_{K0} p_K \left(\frac{N_K}{N} \right) + 2 g_{Na0} x}, \tag{A.5}$$

where c is the factor of proportionality. Because $\left(\frac{N_K}{N} \right)$ is constant, we can separate $e(x, N)$ into the variables N and x:

$$e(x, N) = N \tilde{e}(x). \tag{A.6}$$

The energy function $\tilde{e}(x)$ can be written as

$$\tilde{e}(x) = C \frac{ABx}{Ax + B}, \tag{A.7}$$

with the constants $A = 2 g_{Na0}$, $B = 3 g_{K0} p_K (\frac{N_K}{N})$, and $C = c(E_{Na} - E_K)/6$. Because we define units of energy in this study such that $\tilde{e}(1) = e_{ch} = 1$, the rescaled energy function that is implemented in the Blahut-Arimoto algorithm reads

$$\tilde{e}(x) = \frac{(A + B)x}{Ax + B}. \tag{A.8}$$

It does not depend on the values of the reversal potentials E_{Na} and E_K.

A.2 Numerical Validation. The signaling system with stochastic units operates under additive binomial noise (caused by the random activation and deactivation of the units), whose variance depends on the input x (see equation 2.2). In order to validate the approximation of the information transfer in such a model by equation 4.2, that, strictly speaking, applies only to systems with additive gaussian noise of a fixed, input-independent variance, we performed numerical tests. To this end, the input x was discretized into 1000 equispaced values between zero and one. The information transfer was then calculated according to equation 3.1. The noise distribution $p(k|x)$ was assumed to be binomial as in equation 2.1, and the output distribution $p_k(k)$ was calculated using $p_k(k) = \sum_x p(k|x)p_x(x)$. We found that the information transfer in the stochastic unit system is well approximated by equation 4.2, as shown in Figure 7. For very broad input distributions (e.g., $\sigma_x = 0.16$), where the approximation error is biggest, the deviations are well below 4% for $N > 100$. The relative error is significantly smaller for narrower distributions. The largest deviation below $N = 100$ occurs for $N = 1$ and is less than 11% for broad input distributions $p_x(x)$. A finer discretization into 10, 000 input values gives similar results.

A.3 Efficiency. The most general approach to treat energy as a constraint is the maximization of $I - sE$, where s is a trade-off parameter between energy and information. The higher s, the tighter is the energy budget. This is a Lagrangian multiplier problem, where for a given, fixed energy, the optimal information transfer I and number of channels N have to be determined. However, both the information transfer I and the metabolic cost E depend on only one variable: the number of channels N. So if the energy is fixed, N and I are fixed too. Thus, each point on the curve depicting energy efficiency as a function of the number of channels N (as shown in Figure 2) corresponds to the solution of the general $I - sE$ optimization problem for a particular trade-off parameter s. From these solutions, we have chosen to pay special attention to the ratio I/E. This approach becomes more obvious for the optimization of input distributions with the Blahut-Arimoto algorithm, where $I - sE$ is optimized explicitly before concentrating on the subset of distributions maximizing the ratio I/E.

A.4 Information Capacity and Blahut-Arimoto Algorithm. The Blahut-Arimoto algorithm maximizes the function

$$C[N; s] = \max_{p_x(\cdot)} \left[I[N; p_x(\cdot)] - sE[N; p_x(\cdot)] \right], \tag{A.9}$$

where s is a trade-off parameter that gives the relative importance of energy minimization versus information maximization. The input distribution $p_x(x)$ needs to be discretized. Given an initial input distribution $p_x^0(x)$, the set of conditional probabilities $p(k|x)$, and the expense for each input
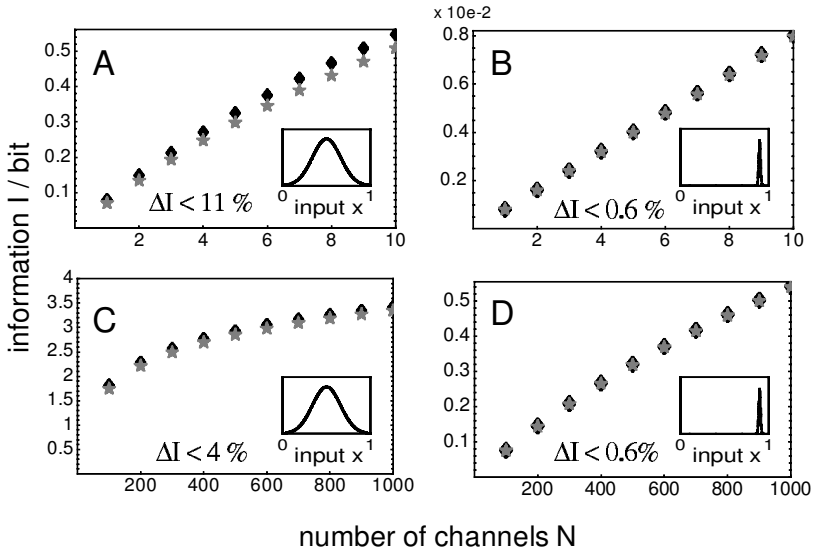
Figure 7: Numerically derived information transfer as a function of N for the model of stochastic units (filled diamonds) in comparison to the information transfer specified by Shannon's formula (gray stars) in equation 4.2. The largest relative difference ΔI is depicted for each set of inputs. The input distributions $p_x(x)$ are shown in the insets. (A) Information transfer for a broad gaussian input distribution $p_x(x)$ with $\sigma_x = 0.16$ and $\overline{x} = 0.5$ for very small numbers of channels N. (B) The same for a narrow input distribution $p_x(x)$ with $\sigma_x = 0.01$ and $\overline{x} = 0.9$. (C) Information transfer for large N and a broad input distribution with parameters as in A. (D) Information transfer for large N and the narrow input distribution used in B. Shannon's formula gives a very good approximation of the information in the channel model. The quality of the approximation increases with rising N and decreasing variance of the input distributions.

symbol, the algorithm iteratively calculates the distribution $p_x(x)$ that maximizes $C[N; s]$. In every step of the iterative algorithm, upper and lower bounds on the information capacity can be derived (Blahut, 1972), which help to estimate the quality of the current probability set.

For the numerical calculations, we discretized the input ($x_j = j/200$ with $j = 0, \ldots, 200$, $x_j \in [0; 1]$). The input distributions maximizing information transfer were calculated for $N = 10, 20, \ldots, 1000$ and 26 values of the parameter s, ranging from 0 (no energy restriction) to 20 (high energy constraint). The cost $\tilde{e}(x)$ was assumed to depend on x according to equation A.8, with $g_{Na0} = 20pS$, $g_{K0} = 20pS$, $p_K = 0.5$, and $N_K/N = 0.5$. The resulting 26 points in C-E space for a given N were subject to a cubic spline interpolation (Figure 5E shows four points and the respective interpolation curve for the

whole set of s; $N = 100$ here.). Afterward, the energy was scaled by multiplication with N according to equation A.6. For the whole study, the fixed cost is always stated in the unit of cost of one open channel (i.e., $b = 500$ corresponds to the cost of 500 open channels). The optimal number of stochastic units, N_{opt}, for a given fixed cost b was determined choosing the N of all calculated N, where the efficiency $(C_N/E)_{max}$ for that b was maximal. Since those values of N are multiples of 10, the data were subsequently fitted with a function f of the form $f(x) = a_1 x^{a_2}$ (a_1 and a_2 being parameters) in order to obtain a smooth graphical representation.

Acknowledgments

References

Abshire, P., & Andreou, A. G. (2000). Relating information capacity to a biophysical model for blowfly photoreceptors. *Neurocomputing, 32*, 9–16.

Abshire, P., & Andreou, A. G. (2001). Capacity and energy cost of information in biological and silicon photoreceptors. *Proceedings of the IEEE, 89*(7), 1052–1064.

Aiello, L. C., & Wheeler, P. (1995). The expensive tissue hypothesis: The brain and the digestive system in human and primate evolution. *Curr. Anthropol., 36*, 199–221.

Alexander, R. M. (1997). A theory of mixed chains applied to safety factors in biological systems. *J. Theor. Biol., 184*, 247–252.

Ames, A. (2000). CNS energy metabolism as related to function. *Brain Res. Rev., 34*, 42–68.

Anderson, J. C., & Laughlin, S. B. (2000). Photoreceptor performance and the coordination of achromatic and chromatic inputs in the fly visual system. *Vision Research, 40*, 13–31.

Arimoto, S. (1972). An algorithm for computing the capacity of an arbitrary discrete memoryless channel. *IEEE Trans. on Info. Theory, IT-18*, 14–20.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signalling in the grey matter of the brain. *J. Cereb. Blood Flow Metab., 21*, 1133–1145.

Baddeley, R., Abbott, L. F., Booth, M. C. A., Sengpiel, F., Freeman, T., Wakeman, E. A., & Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. Lond. B, 264*, 1775–1783.

Balasubramanian, V., Kimber, D., & Berry, M. J. I. (2001). Metabolically efficient information processing. *Neural Comput., 13*, 799–816.

Blahut, R. E. (1972). Computation of channel capacity and rate distortion functions. *IEEE Trans. on Info. Theory, IT-18*, 460–473.

de Polavieja, G. G. (in press). Errors drive the evolution of biological signalling to costly codes. *J. Theor. Biol.*

de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R., and Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science, 275,* 1805–1808.

Gronenberg, W., & Liebig, J. (1999). Smaller brains and optic lobes in reproductive worker of the ant *Harpegnathos*. *Naturwiss., 86,* 343–345.

Kern, M. F. (1985). Metabolic rate and the insect brain in relation to body size and phylogeny. *Comp. Biochem. Physiol., 81A,* 501–506.

Laughlin, S. B. (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch., 36c,* 910–912.

Laughlin, S. B. (1989). The reliability of single neurons and circuit design: A case study. In R. Durbin, C. Miall, & G. Mitchison (Eds.), *The computing neuron* (pp. 322–336). Reading, MA: Addison-Wesley.

Laughlin, S. B. (1994). Matching coding, circuits, cells, and molecules to signals: General principles of retinal design in the fly's eye. *Prog. Ret. Eye Res., 13,* 165–196.

Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Curr. Opin. Neurobiol., 11,* 475–480.

Laughlin, S. B., Anderson, J. C., O'Carroll, D. C., & de Ruyter van Steveninck, R. R. (2000). Coding efficiency and the metabolic cost of sensory and neural information. In R. Baddeley, P. Hancock, & P. Foldiak (Eds.), *Information theory and the brain* (pp. 41–61). Cambridge: Cambridge University Press.

Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. *Nat. Neurosci., 1*(1), 36–41.

Levy, W. B., & Baxter, R. A. (1996). Energy efficient neural codes. *Neural Comput., 8*(3), 531–543.

Lutz, P., & Nilsson, G. E. (1994). *The brain without oxygen.* Austin: R. G Landes.

Martin, R. D. (1996). Scaling of the mammalian brain—the maternal energy hypothesis. *News in Physiol. Sci., 11,* 149–156.

Nilsson, G. E. (1996). Brain and body oxygen requirements of *Gnathonemus petersii,* a fish with an exceptionally large brain. *J. Exp. Biol., 199,* 603–607.

Rolfe, D. F. S., & Brown, G. C. (1997). Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol. Rev., 77,* 731–758.

Sarpeshkar, R. (1998). Analog versus digital: Extrapolating from electronics to neurobiology. *Neural Comput., 10,* 1601–1638.

Schreiber, S. (2000). *Influence of channel noise and metabolic cost on neural information transmission.* Unpublished diploma thesis available at the library of Humboldt-University Berlin, Germany.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Siesjö, B. (1978). *Brain energy metabolism.* New York: Wiley.

White, J. A., Rubinstein, J. T., & Kay, A. R. (2000). Channel noise in neurons. *TINS, 23*(3), 131–137.

Zador, A. (1998). Impact of synaptic unreliability on the information transmitted by spiking neurons. *J. Neurophysiol., 79,* 1219–1229.