EFFICIENT INFORMATION TRANSFER BY POISSON NEURONS

LUBOMIR KOSTAL

Institute of Physiology of the Czech Academy of Sciences
Videnska 1083, 14220 Prague 4, Czech Republic

SHIGERU SHINOMOTO

Department of Physics, Kyoto University
Sakyo-ku, Kyoto, 606-8502, Japan

ABSTRACT. Recently, it has been suggested that certain neurons with Poissonian spiking statistics may communicate by discontinuously switching between two levels of firing intensity. Such a situation resembles in many ways the optimal information transmission protocol for the continuous-time Poisson channel known from information theory. In this contribution we employ the classical information-theoretic results to analyze the efficiency of such a transmission from different perspectives, emphasising the neurobiological viewpoint. We address both the ultimate limits, in terms of the information capacity under metabolic cost constraints, and the achievable bounds on performance at rates below capacity with fixed decoding error probability. In doing so we discuss optimal values of experimentally measurable quantities that can be compared with the actual neuronal recordings in a future effort.

1. **Introduction.** The problem of information processing and transmission in the brain is historically one of the most intensively studied topics in neurosciences [2, 40, 10]. Frequently, the theoretical approach to this problem relies on the methods of information theory [15] with emphasis on channel capacity as the ultimate fidelity criterion [42, 33, 11, 36]. The operational interpretation of channel capacity relies on an essentially digital transmission protocol (not necessarily binary) and a special encoding-decoding setup known as the separation assumption [15, 43]. Additionally, the information rate equal to the capacity is only an asymptotic quantity, assuming arbitrarily reliable communication, infinite decoder complexity and delay. The unconstrained (or asymptotically achievable) amount of information, however, might not be the main objective for biological systems, as metabolic costs and real-time information processing should definitely be considered [45, 30, 46]. In this paper we attempt to include multiple factors that might affect the efficiency of the actual information transmission, thus continuing the effort started in [28, 30, 29].

The available information-theoretic methods do not allow to include many of the detailed biological phenomena easily [15, 44]. In particular, different time scales of the dynamics dictate that dependence on the history and feedback is frequently present in neural systems [5]. On the other hand, simplified models often provide

---

a satisfactory description of experimental data [51, 16, 26]. The classic Poissonian assumption of neural spiking statistics seems sufficient under broad circumstances [39, 48] and the capacity of Poisson neurons has been discussed in the neuroscientific context too [22, 23]. The key advantage here is that a plethora of results is available for the Poisson communication [24, 9, 54, 55, 37], so that the model can be employed also outside of the neurobiological context (e.g., in [50]).

It was observed in [38] that certain neurons with Poissonian spiking statistics may communicate by discontinuously switching between two discrete levels of firing intensity. We calculate the asymptotic limits on reliable communication under such circumstances and we address the problem of the *effective* information rate, provided that the coding system is constrained in its bandwidth and complexity. We believe that the bandwidth limitation might be of special interest since neural systems are generally limited by the finite speed of the underlying chemical and electric processes [10]. The main goal of this paper is therefore to provide an information-theoretic interpretation of the results recently obtained by Mochizuki and Shinomoto [38]. We are convinced that our effort is timely as the problem of constrained information transmission has been attracting attention in the theoretical neuroscience community recently [49, 45, 30, 46, 29].

2. Methods.

2.1. **Information capacity of the Poisson neuron.** Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and $\mathcal{F}_t$ a filtration on $\mathcal{F}$. The neuronal input is described by a stationary stochastic process $\lambda_t \geq 0, t \in [0, T]$ adapted to $\mathcal{F}_t$. The neuronal response is given by the doubly stochastic Poisson process N_t with intensity $\lambda_t + \mu_0$. In this scenario we assume that λ_t is proportional to the driving synaptic current [25, 16, 31], $\mu_0 \geq 0$ is the spontaneous activity rate and N_t is the number of postsynaptic spikes observed up to time t. The following two constraints are imposed on the input signal. The peak-amplitude constraint,

$$0 \leq \lambda_t \leq L, \tag{1}$$

and the average power constraint

$$\mathbb{E}_\lambda \int_0^T \lambda_t \, \mathrm{d}t \leq \varrho LT, \tag{2}$$

where $0 \leq \varrho \leq 1$ is the maximum allowed average-to-peak ratio of the input signal.

The information capacity (in nats per second) of the point process N_t is given by

$$C = \lim_{T \to \infty} \sup_\lambda \frac{1}{T} I_T(\lambda, N), \tag{3}$$

where $\sup_\lambda$ is over all signals λ_t satisfying Eqs. (1) and (2), with no other restrictions on the form of the trajectories or their statistics. The mutual information between λ_t and N_t is

$$I_T(\lambda, N) = \mathbb{E}_\lambda \int_0^T \left[\varphi(\lambda_t) - \varphi(\hat{\lambda}_t) \right] \mathrm{d}t, \tag{4}$$

where

$$\varphi(x) = (\mu_0 + x) \log(\mu_0 + x) - \mu_0 \log \mu_0, \tag{5}$$

$\hat{\lambda}_t = \mathbb{E}_\lambda(\lambda_t | \mathcal{F}_t^N)$ and $\mathcal{F}_t^N$ is the canonical filtration of N_t [6, Ch. 6.5]. Hence, the conditional mean $\hat{\lambda}_t$ is the minimum mean-square error estimate of λ_t based on the

history of the process N_t. It was shown in [9] (see also [24] and [6]) that, remarkably, Eq. (3) can be evaluated in a closed form,

$$C(\varrho) = L\big[(1+s)r\log(1+s) + (1-r)s\log s - (r+s)\log(r+s)\big], \qquad (6)$$

where

$$s = \mu_0/L, \qquad (7)$$

$$r = \min(\varrho, \xi), \qquad (8)$$

$$\xi = \frac{1}{e}\frac{(1+s)^{1+s}}{s^s} - s. \qquad (9)$$

If the neuron is not spontaneously active, $\mu_0 = 0$, then Eq. (6) reduces considerably to $C(\varrho) = -L\min(\varrho, e^{-1})\log[\min(\varrho, e^{-1})]$.

The capacity-achieving input signal λ_t^* is a limiting case of the random telegraph wave, i.e., of the stationary process with a piecewise constant path taking only the extremal values 0 and L. More precisely, let n be a positive integer and consider the process λ_t^n which admits the representation [9, 6]

$$\lambda_t^n = \lambda_0^n + n\int_0^t 1(\lambda_s^n = 0)\,\mathrm{d}s - \frac{n(1-r)}{r}\int_0^t 1(\lambda_s^n = L)\,\mathrm{d}s + m_t^n, \qquad (10)$$

where r is given by Eq. (8), $P[\mu_0^n = L] = r$ and m_t^n is a martingale. The capacity achiever is the process $\lambda_t^* = \lim_{n\to\infty}\lambda_t^n$. As n grows the rate of transitions between states $\lambda^n = 0$ and $\lambda^n = L$ is unbounded and therefore the optimal input has infinite bandwidth. The optimal average-to-peak ratio of the input signal equals r. Note that if $\mu_0 = 0$ and $\varrho = 1$ (no average-power constraint on λ_t) then the probability that $\lambda_t = L$ at any given time equals e^{-1}.

2.2. **Metabolic cost of neural activity.** Both empirical and theoretical studies suggest that the metabolic cost of neuronal spiking activity, in terms of ATP molecules (ATPm) expenditure, is proportional to the firing rate [4, 3]. By employing the linear model of Laughlin and Attwell [3] we naturally define the average metabolic cost W (in ATPm per second) evoked by the optimal input signal as

$$W = (\mu_0 + rL)\kappa + \beta, \qquad (11)$$

where $\kappa = 0.71 \times 10^9$ ATPm is the cost of a single spike and $\beta = 0.34 \times 10^9$ ATPm/s is the basic metabolic rate required to maintain the membrane potential. In what follows we express $C(\varrho)$ in terms of W as $C(W)$, even though due to Eq. (8) the same value of W may correspond to different ϱ for high enough ϱ.

We propose the information efficiency, E (in bit per ATPm), as the ratio of information capacity to the associated metabolic cost,

$$E = \frac{C(W)}{W}. \qquad (12)$$

The maximum efficiency with respect to the cost, $E^* = \max_W E$ (also known as the capacity per unit cost [52]), represents the maximum amount of information per ATPm that can be communicated at arbitrarily low probability of channel decoding error. The rationale for adopting Eq. (12) as the efficiency measure lies in the crucial importance of balancing the performance versus the metabolic workload for living organisms [18, 34, 30].

2.3. **Coding capacity and Wyner's code.** The extension of Shannon's discrete-time channel coding theorem for general continuous-time channels is mathematically non-trivial [41, 20, 53]. However, in his seminal work Wyner [54, 55] proved (among other things) the achievability and converse theorems for the information capacity in Eq. (6) by employing the equivalence between the Poisson continuous-time channel and a certain discrete-time memoryless channel. In other words the information capacity of the Poisson neuron equals its coding capacity. Hence, for any information rate up to capacity, $0 \leq R < C$, and arbitrary $\varepsilon > 0$ there exists a channel code with $M \geq e^{RT}$ codewords (for T sufficiently large) such that the average probability of decoding error P_e satisfies $P_e \leq \varepsilon$ [15].

The construction of Wyner's code ensemble with integer parameters $M > 1$ and $k \geq 1$,

$$M = \lfloor e^{RT} \rfloor, \tag{13}$$

$$k = \lfloor rM \rfloor, \tag{14}$$

proceeds by dividing the interval $[0, T]$ into $\binom{M}{k}$ subintervals and by assigning the value $\lambda = L$ into $(k/M)\binom{M}{k}$ bins and $\lambda = 0$ to the remaining bins (see [54, Sec. III.A] for details). In this way, M distinct input signal waveforms (codewords) are created, each satisfying Eq. (2) while bearing similarity to the process λ_t^n in Eq. (10). The maximum likelihood decoder observes the process N_t and identifies the input waveform based on the spike counts. The probability of mismatch (decoding error) for each input waveform is $P_{e,m}$. The sequence of codes provided by Wyner is essentially optimal, satisfying the Gallager's random coding bound [15]. Therefore we can define the achievable average probability of decoding error at rate R as

$$P_e \triangleq e^{-E(R)T}, \tag{15}$$

where $E(R) > 0$ for all $0 \leq R < C$ and $P_e = \sum_{m=1}^{M} P_{e,m}/M$. The coordinate $\{R, E(R)\}$ can be expressed in an implicit parametric form in terms of $p \in (0, 1]$ following [54] as

$$E(R) = \begin{cases} R_0 - R, & 0 \leq R < R(1), \\ LE_1(p) - pR, & R(1) \leq R < C, \end{cases} \tag{16}$$

with the following functions defined,

$$\tau(p) = \left(1 + \frac{1}{s}\right)^{1/(1+p)} - 1, \tag{17}$$

$$q(p) = \min\left\{\varrho, \frac{1}{\tau(p)}\left[\left(\frac{1}{(1+p)s\tau(p)}\right)^{1/p} - 1\right]\right\}, \tag{18}$$

$$R \equiv R(p) = Ls\left[\frac{[1 + q(p)\tau(p)]^p}{1+p}q(p)\left(1 + \frac{1}{s}\right)^{1/(1+p)}\log\left(1 + \frac{1}{s}\right) - \right.$$
$$\left. - [1 + q(p)\tau(p)]^{1+p}\log\left(1 + q(p)\tau(p)\right)\right], \tag{19}$$

$$E_1(p) = q(p) + s - [1 + \tau(p)q(p)]^{1+p}s, \tag{20}$$

and s given by Eq. (7). The value $R_0 = E(0)$ is the cutoff rate [35],

$$R_0 = Lq^*(1 - q^*)\left(\sqrt{1+s} - \sqrt{s}\right)^2, \tag{21}$$

$$q^* = \min(\varrho, 1/2). \tag{22}$$

The time window T thus plays a similar role to the channel code blocklength in discrete-time channels [15, Theorem 7.3.2].

2.4. **Effective information rate and required bandwidth.** The capacity represents the maximal information rate achievable under the stringent reliability criterion $P_e \to 0$. Approaching capacity then requires unbounded coding windows T and infinitely fast switching rates of the input signal [9, 54]. It is practically impossible to implement such information transmission schemes. The error exponent $E(R)$ in Eq. (15) relates the information rate R, coding window T and the probability of decoding error P_e. Therefore, we employ it to analyze the possibility of effective information transfer at rates below capacity and tolerable value of P_e as follows.

For a fixed value of P_e and rate R we define the required length of coding window T as

$$T = \max(T_P, T_k), \tag{23}$$

where

$$T_P = -\frac{\log P_e}{E(R)}, \tag{24}$$

$$T_k = \frac{\log \lceil 1/r \rceil}{R}. \tag{25}$$

Then Eq. (24) follows directly from Eq. (15) and guarantees that the error probability does not exceed P_e. The condition in Eq. (25) is necessary since $k \geq 1$ in Wyner's code, and follows from the combination of Eq. (14) for $k = 1$ and Eq. (13). Once the coding window T is determined, the M, k parameters of Wyner's code may be calculated through Eqs. (13) and (14). The maximal required bandwidth B of the input signal is of particular practical importance. Regularly, the bandwidth of a unit pulse is defined to be equal to its inverse duration in engineering applications [12], hence from the properties of Wyner's code we have

$$B \triangleq \binom{M}{k} \frac{1}{T}. \tag{26}$$

In other words, Eq. (26) implicitly defines the *effective* information rate $R_{\text{eff}}(B)$, which is achievable with input signal bandwidth not exceeding B at some fixed error probability P_e.

3. **Results.** In the following we evaluate the information rates and capacities below in bits, using the standard conversion 1 "bit" $= \log(2)$ "nat".

First we examine the dependence of the neuron capacity given by Eq. (6) on the metabolic expenditure given by Eq. (11), more precisely we examine the capacity-cost function $C(W)$. We set the peak input firing rate to $L = 50$ Hz and investigate the effect of the spontaneous activity of four different intensities $\mu_0 = \{0.1, 1, 5, 10\}$ Hz. The chosen values are entirely physiological (note that the output firing rate peaks at $L + \mu_0$ Hz) and represent a rather generic situation under various experimental conditions (e.g., [38, 13, 17]). We do not consider peak firing rates $L > 50$ Hz in order to guarantee the validity of the linear cost model in Eq. (11) [3] and of the Poissonian approximation to the real neuronal firing activity [21]. In practice it is possible to determine L and μ_0 by fitting the hidden Markov model to the given the experimentally observed neuronal activity, as has been done in [38]. Alternatively, one may fit the hidden Markov model to the data by fixing two firing

rates in advance. Here the two firing rates may be chosen by hand or using other principles.

The respective capacity-cost functions are shown in Fig. 1a. The metabolic cost W for each μ_0 spans the whole possible range of values since the parameter ϱ in Eq. (2) varies continuously in the interval $(0, 1)$. The $C(W)$ departs from zero at the minimum possible cost, $W_{\min} = \mu_0 \kappa + \beta$, which corresponds to the metabolic cost of the spontaneous firing rate μ_0 and the baseline cost, see Eq. (11). The unconstrained channel capacity is the limit of $C(W)$ as the metabolic cost is allowed to grow. Under all four spontaneous activity levels the unconstrained capacity exceeds 10 bit/s. The efficiency E defined by Eq. (12) balances the information capacity and the required metabolic expenditure (Fig. 1b). Note that the factor $L/\mu_0 = 1/s$ essentially describes the signal-to-noise ratio (SNR) of the Poisson neuron. The capacity in Eq. (6), as a function of SNR, is proportional to L and hence we may obtain different capacity values for situations with equal SNR. We see that as the SNR increases, the point of maximal efficiency E^* moves towards lower costs. (At the same time the cost of 1 bit (equal to $1/E^*$) decreases, as expected intuitively.) The point of maximal efficiency is less pronounced for low SNR.

An alternative point of view is provided by analyzing the best possible performance when the metabolic cost is fixed, $W = \text{const}$. It follows from Eq. (11) that W and the average postsynaptic firing rate $\langle f \rangle$ are linearly related due to $\langle f \rangle = \mu_0 + rL$. We therefore maximize the capacity as a function of s, L along the curve $\langle f \rangle = (s + r)L = \text{const}$. (note that r is a function of s). Since L acts only as a scaling factor in Eq. (6) it is possible to verify that the solution to this constrained maximization problem requires $s = 0$. In other words, for the given value of $\langle f \rangle$, the Poisson neuron with parameters $\mu_0 = 0$ and $L = \langle f \rangle / \min(\varrho, 1/e)$ attains the largest capacity.

The effective rate for $B = 2\,\text{Hz}$ at $P_e = 10^{-5}$ is shown in Fig. 1c. For convenience we neglect that M and k are integers and set $M = e^{RT}$, $k = rM$ together with

$$\binom{M}{k} = \frac{\Gamma(M + 1)}{\Gamma(k + 1)\Gamma(M - k + 1)}, \tag{27}$$

where $\Gamma(x)$ is the gamma function [1]. The most striking difference with respect to Fig. 1a is the significant drop in information transfer, approximately one order of magnitude large. In addition, if the metabolic cost is neglected, the difference between maximal R_{eff} becomes negligible with growing SNR, e.g., there is virtually no difference between $\mu_0 = 0.1\,\text{Hz}$ and $\mu_0 = 1\,\text{Hz}$. Similarly to Fig. 1b the ratio R_{eff}/W corresponds to the achievable efficiency (Fig. 1d). The notable difference here, besides the respective values of the maxima, is that the points of maximal efficiency occur at lower values of the metabolic cost (shown by vertical lines).

An alternative point of view is provided by fixing the metabolic cost and varying the maximum allowed input signal bandwidth. Although currently do not have enough information on the biologically valid range of B, we believe that relatively small values, $B < 10\,\text{Hz}$, are plausible. Fig. 2a shows such a situation for $\varrho = 0.2$. The value $\varrho = 0.2$ was chosen for illustration only, there is a qualitative correspondence between the results for different ϱ (not shown). In terms of the metabolic cost and the capacity we have $W = 8.6\,\text{ATPm/s}$ and $C(W) = 22.6\,\text{bit/s}$ for $\mu_0 = 0.1\,\text{Hz}$, $W = 9.2\,\text{ATPm/s}$ and $C(W) = 19.8\,\text{bit/s}$ for $\mu_0 = 1\,\text{Hz}$, $W = 12.0\,\text{ATPm/s}$ and $C(W) = 14.3\,\text{bit/s}$ for $\mu_0 = 5\,\text{Hz}$, $W = 15.6\,\text{ATPm/s}$ and $C(W) = 11.0\,\text{bit/s}$ for $\mu_0 = 10\,\text{Hz}$. At this value of the input signal average-to-peak ratio the growth of
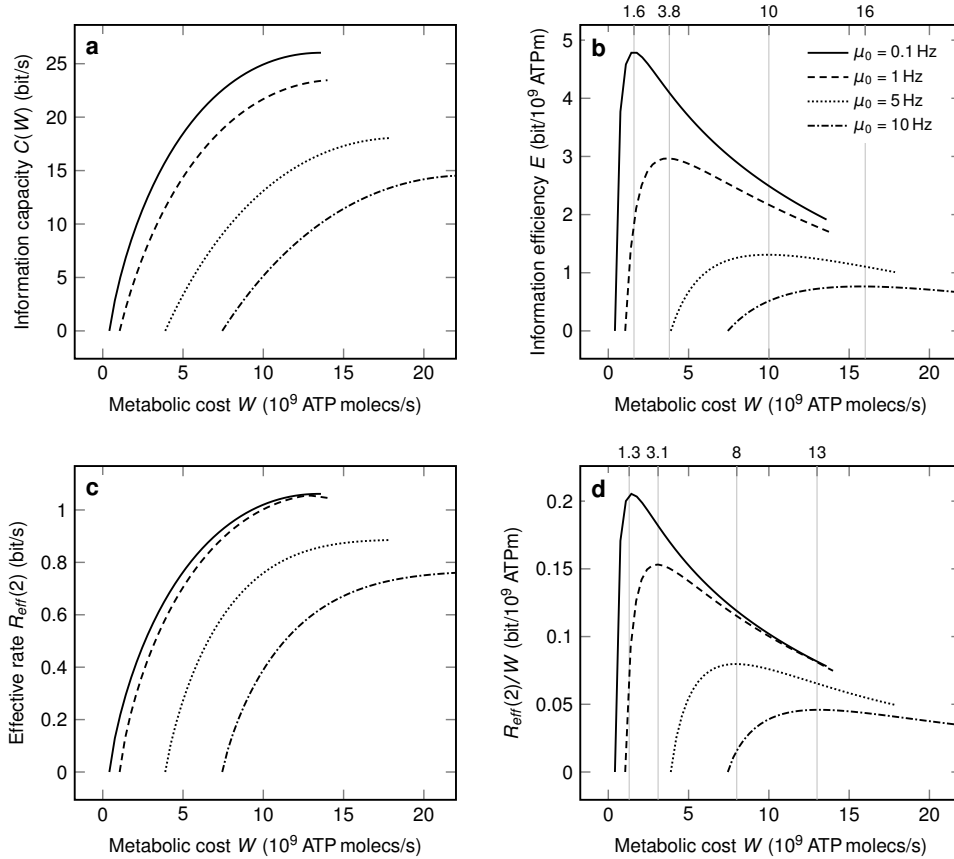
FIGURE 1. Capacity, efficiency and the effective information transfer rate of the Poisson neuron in dependence on the average (post-synaptic) metabolic cost W. Four different values of the spontaneous firing rate μ_0 are examined, the input signal (presynaptic firing intensity) is restricted to $\lambda_t \in [0, L]$ and $L = 50\,\mathrm{Hz}$. The capacity-cost function (**a**) exhibits the law of diminishing returns as W grows. The efficiency $C(W)/W$ (**b**) shows a pronounced optimum for higher signal-to-noise (SNR) ratios L/μ_0. Both the capacity and the efficiency are asymptotic quantities in terms of coding-decoding delay and complexity. The effective information rate R_{eff} (**c**) is a non-asymptotic quantity with explicit coding considerations, i.e., λ_t bandwidth not exceeding $2\,\mathrm{Hz}$ and decoding error probability $P_e = 10^{-5}$. The maximal R_{eff} is only about $5\,\%$ of the unconstrained capacity value and the benefit of increased SNR between $\mu_0 = 1\,\mathrm{Hz}$ and $\mu_0 = 0.1\,\mathrm{Hz}$ is negligible. The balance between the metabolic cost and the effective rate (**d**) also achieves only $5\,\%$ of the asymptotic values in (b), in addition, the optima occur for smaller metabolic cost (vertical lines).
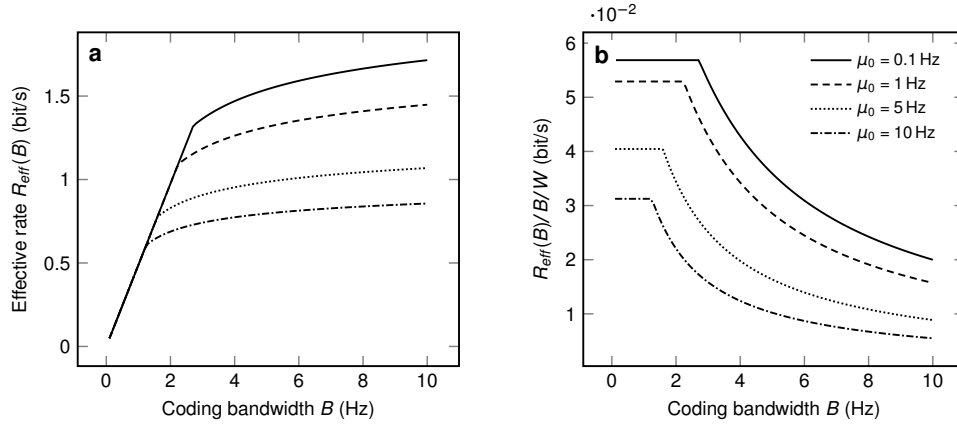
FIGURE 2. Effective information rate R_{eff} when the input signal
bandwidth does not exceed B under constant $\varrho = 0.2$, see Eq. (2).
As in Fig. 1c, d it holds $L = 50\,\text{Hz}$ and $P_e = 10^{-5}$. For small band-
widths the effective rate grows linearly and independently of the
SNR (level of spontaneous activity μ_0) up to the critical bandwidth
$B(\mu_0)$ (**a**). The information rate per metabolic cost per allocated
bandwidth (**b**) considers the three most important aspects of prac-
tical information transmission jointly. The rapid drop in $R_{\text{eff}}/W/B$
after the critical bandwidth shows that increasing the transmission
rate is not equally easy in terms of practical coding parameters.

R_{eff} with B is initially linear and independent of μ_0. This trend continues up to
certain critical B (dependent on μ_0), from which on the growth of information rate
becomes much slower. The ratio $R_{\text{eff}}/B/W$ (Fig. 2b) balances the information rate
vs. the allocated signal bandwidth *vs.* the incurred metabolic expenses, although
the actual importance of each factor can hardly be determined unquestionably.
Nonetheless the key observation here is the significant drop in information-transfer
capabilities above the critical bandwidth (which grows with SNR).

Finally, we illustrate behavior of the key Wyner's code parameters in dependence
on the transmission rate R at bandwidth not exceeding $B = 2\,\text{Hz}$ and $\varrho = 0.2$. The
required coding window T is shown in Fig. 3a. The actual value of T is analogous
to the discrete-time channel code blocklength and hence the decoding delay and
complexity increase with T [15, 7, 29]. We observe that for a fixed constraint ϱ and
maximal allowed bandwidth B there exists an information transmission rate $R > 0$
achieving smallest T. The existence of such an "optimal" R stems from the condition
in Eq. (23). If the value of P_e was decreased, the minimum of T would become
more pronounced and occur at smaller rates (not shown). The coding bandwidth in
Eq. (26) combines the coding window duration with the number of time divisions
$\binom{M}{k}$. The number of these divisions grows explosively once the optimal rate in
Fig. 3a is crossed (Fig. 3b). Wyner's coding waveforms thus approach the capacity
achieving input signal in Eq. (10) rapidly.

4. **Discussion and conclusions.** In this contribution we have analyzed the infor-
mation coding and transmission efficiency of a Poissonian neuron. We have deter-
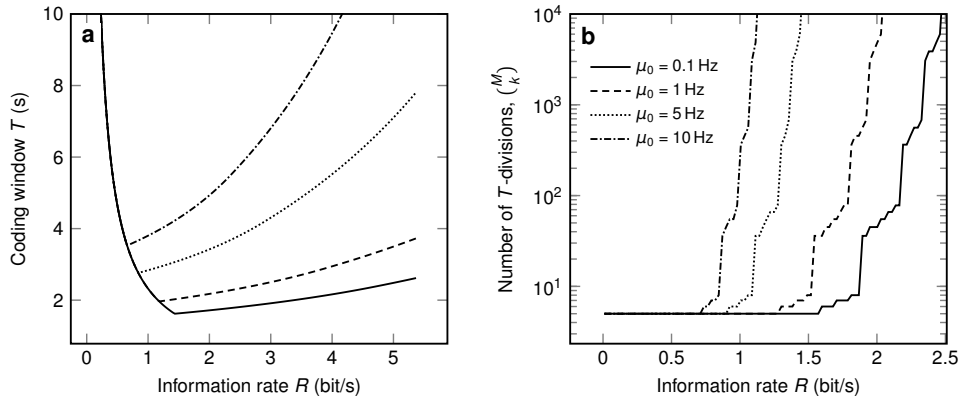mined the channel capacity and efficiency (Fig. 1a, b) and stressed their asymptotic

FIGURE 3. Detailed parameters of Wyner's code with $B \leq 2\,\mathrm{Hz}$ and $\varrho = 0.2$ in dependence on the transmission rate ($L = 50\,\mathrm{Hz}$ and $P_e = 10^{-5}$). The required coding window duration T corresponds to the classical discrete-time channel code blocklength. The decoding complexity and delay grows with T so it is desirable to minimize T for the given rate R and P_e (**a**). For all investigated SNRs there exists an information rate with minimal T and this rate increases with the SNR. The coding window $[0, T]$ is divided into $\binom{M}{k}$ sub-intervals. The number of these bins grows explosively once the optimal information rate (minimal T) is crossed (**b**). Consequently, the required bandwidth of the input signal grows too (Fig. 2a).

properties and the impossibility of actually achieving these rates practically. The focus of our effort is the explicit consideration of decoding error probability and its influence on the code complexity (bandwidth or duration). These considerations might be of crucial importance for finite-sized, power-constrained and real-time operating communication systems – such as neurons. To our best knowledge, though, this part of information (or communication) theory has not been applied to neuroscience, with the exception of [29]. We have found that once the balance between information, metabolic cost and code parameters is taken into account, the achievable information rates drop significantly with respect to their asymptotic counterparts (Fig. 1c, d). Similar conclusions were reached in [29] for the case of a realistic Hodgkin-Huxley neuronal model. In addition, the progress towards the higher information rates is accompanied by sudden and rapid increase in certain undesirable code parameters (Figs. 2 and 3). The difficulties associated with crossing the cutoff rate in Eq. (21) are well known [35, 37]. Here we have demonstrated similar effects at lower rates with respect to parameters of potential neurobiological importance.

Throughout this paper it has been assumed that the input signal (the presynaptic firing rate λ_t) has a very restricted form, taking only two possible values. This assumption stems directly from the mathematical fact that such a signal is optimal (asymptotically) for the Poisson communication, and from the recent analysis of Mochizuki and Shinomoto [38] of real neuron behavior. However, our methods do not answer the question whether, under certain coding constraints, a continuously

varied λ_t might be beneficial. In this context, Kostal and Kobayashi [29] have shown, without imposing the discreteness constraint on the presynaptic firing rate, that coding complexity restrictions result in an input with very few and well separated levels. In addition, the binary input signal usually near-achieves the information capacity values in systems with low SNR [27, 28, 19].

Finally, few remarks about the interpretation of our results are in place. First, the definition of achievable P_e in Eq. (15) is only approximate for small T. The reason lies in the fact that unlike the unconstrained Gallager's bound which holds for all code lengths [15, Theorem 5.6.2], the power constraint in Eq. (10) prohibits an equally elegant formulation [15, Theorem 7.3.2]. On the other hand, the Gallager's bound is the *upper* bound on the smallest possible P_e, so it is entirely possible that a code with P_e satisfying Eq. (15) exists even for very small T. Second, the capacity in Eq. (6) cannot be further increased by causal feedback, that is, λ_t may depend on $N_\tau, \tau \in [0, t]$ [6]. The feedback however allows for the existence of simpler codes at given $R < C$ and P_e, and therefore increases the value of the error exponent [32]. The capacity thus cannot be further increased by a non-Poissonian neuron, provided that the presynaptic input still satisfies the constraints in Eq. (1) and (2) and provided that the spontaneous activity is Poissonian with intensity μ_0. For further capacity-related results on the Poisson channel subject to additional constraints see [14, 8, 47].

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, Dover, New York, 1992.

[2] E. D. Adrian, *Basis of Sensation*, W. W. Norton and Co., New York, 1928.

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

[4] V. Balasubramanian and M. J. Berry, A test of metabolically efficient coding in the retina, *Netw. Comput. Neural Syst.*, **13** (2002), 531–552.

[5] W. Bialek, *Biophysics: Searching for Principles*, Princeton University Press, Princeton, 2012.

[6] P. Brémaud, *Point Processes and Queues, Martingale Dynamics*, Springer, New York, USA, 1981.

[7] G. Caire, S. Shamai and S. Verdu, Noiseless data compression with low-density parity-check codes, in *Advances in Network Information Theory* (eds. P. Gupta, G. Kramer and A. J. van Wijngaarden), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 66, American Mathematical Society, 2004, 263–284.

[8] K. Chakraborty and P. Naryan, The poisson fading channel, *IEEE Trans. Inf. Theory*, **53** (2007), 2349–2364.

[9] M. Davis, Capacity and cutoff rate for Poisson-type channels, *IEEE Trans. Inf. Theory*, **26** (1980), 710–715.

[10] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, MIT Press, 2001.

[11] R. R. de Ruyter van Steveninck and S. B. Laughlin, The rate of information transfer at graded-potential synapses, *Nature*, **379** (1996), 642–644.

[12] R. C. Dorf, *The Electrical Engineering Handbook*, CRC Press, Boca Raton, USA, 1997.

[13] P. Duchamp-Viret, L. Kostal, M. Chaput, P. Lansky and J.-P. Rospars, Patterns of spontaneous activity in single rat olfactory receptor neurons are different in normally breathing and tracheotomized animals, *J. Neurobiol.*, **65** (2005), 97–114.

[14] M. R. Frey, Information capacity of the poisson channel, *IEEE Trans. Inf. Theory*, **37** (1991), 244–256.

[15] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Sons, Inc., New York, USA, 1972.

[16] W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge University Press, Cambridge, 2002.

[17] A. Gremiaux, T. Nowotny, D. Martinez, P. Lucas and J.-P. Rospars, Modelling the signal delivered by a population of first-order neurons in a moth olfactory system, *Brain Res.*, **1434** (2012), 123–135.

[18] A. Hasenstaub, S. Otte, E. Callaway and T. J. Sejnowski, Metabolic cost as a unifying principle governing neuronal biophysics, *Proc. Natl. Acad. Sci. U.S.A.*, **107** (2010), 12329–12334.

[19] J. Huang and S. P. Meyn, Characterization and computation of optimal distributions for channel coding, *IEEE Trans. Inf. Theory*, **51** (2005), 2336–2351.

[20] S. Ihara, *Information Theory for Continuous Systems*, World Scientific Publishing, Singapore, 1993.

[21] E. Javel and N. F. Viemeister, Stochastic properties of cat auditory nerve responses to electric and acoustic stimuli and application to intensity discrimination, *J. Acoust. Soc. Am.*, **107** (2000), 908–921.

[22] D. H. Johnson, Point process models of single-neuron discharges, *J. Comput. Neurosci.*, **3** (1996), 275–299.

[23] D. H. Johnson and I. N. Goodman, Inferring the capacity of the vector Poisson channel with a Bernoulli model, *Netw. Comput. Neural Syst.*, **19** (2008), 13–33.

[24] Y. M. Kabanov, The capacity of a channel of the poisson type, *Theor. Prob. App.*, **23** (1978), 143–147.

[25] R. Kobayashi, S. Shinomoto and P. Lansky, Estimation of time-dependent input from neuronal membrane potential, *Neural Comput.*, **23** (2011), 3070–3093.

[26] R. Kobayashi, Y. Tsubo and S. Shinomoto, Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold, *Front. Comput. Neurosci.*, **3** (2009), p9.

[27] L. Kostal, Information capacity in the weak-signal approximation, *Phys. Rev. E*, **82** (2010), 026115.

[28] L. Kostal, Approximate information capacity of the perfect integrate-and-fire neuron using the temporal code, *Brain Res.*, **1434** (2012), 136–141.

[29] L. Kostal and R. Kobayashi, Optimal decoding and information transmission in Hodgkin-Huxley neurons under metabolic cost constraints, *Biosystems*, **136** (2015), 3–10.

[30] L. Kostal, P. Lansky and M. D. McDonnell, Metabolic cost of neuronal information in an empirical stimulus-response model, *Biol. Cybern.*, **107** (2013), 355–365.

[31] S. Koyama and L. Kostal, The effect of interspike interval statistics on the information gain under the rate coding hypothesis, *Math. Biosci. Eng.*, **11** (2014), 63–80.

[32] A. Lapidoth, On the reliability function of the ideal poisson channel with noiseless feedback, *IEEE Trans. Inf. Theory*, **39** (1993), 491–503.

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

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

[35] J. L. Massey, Capacity, cutoff rate, and coding for a direct-detection optical channel, *IEEE Trans. Commun.*, **29** (1981), 1615–1621.

[36] M. D. McDonnell, S. Ikeda and J. H. Manton, An introductory review of information theory in the context of computational neuroscience, *Biol. Cybern.*, **105** (2011), 55–70.

[37] R. J. McEliece, Practical codes for photon communication, *IEEE Trans. Inf. Theory*, **27** (1981), 393–398.

[38] Y. Mochizuki and S. Shinomoto, Analog and digital codes in the brain, *Phys. Rev. E*, **89** (2014), 022705.

[39] R. Moreno-Bote, Poisson-like spiking in circuits with probabilistic synapses, *PLoS Comput. Biol.*, **10** (2014), e1003522.

[40] D. H. Perkel and T. H. Bullock, Neural coding, *Neurosci. Res. Prog. Sum.*, **3** (1968), 405–527.

[41] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, Inc., San Francisco, 1964.

[42] F. Rieke, R. de Ruyter van Steveninck, D. Warland and W. Bialek, *Spikes: Exploring the Neural Code*, MIT Press, Cambridge, 1999.

[43] B. Rimoldi, Beyond the separation principle: A broader approach to source-channel coding, in *4th Int. ITG Conf.*, VDE Verlag, Berlin, 2002, 233–238.

[44] P. Sadeghi, P. O. Vontobel and R. Shams, Optimization of information rate upper and lower bounds for channels with memory, *IEEE Trans. Inf. Theory*, **55** (2009), 663–688.

[45] B. Sengupta, S. B. Laughlin and J. E. Niven, Balanced excitatory and inhibitory synaptic currents promote efficient coding and metabolic efficiency, *PLoS Comput. Biol.*, **9** (2013), e1003263.

[46] B. S. Sengupta and M. B. Stemmler, Power consumption during neuronal computation, *Proc. IEEE*, **102** (2014), 738–750.

[47] S. Shamai and A. Lapidoth, Bounds on the capacity of a spectrally constrained poisson channel, *IEEE Trans. Inf. Theory*, **39** (1993), 19–29.

[48] W. R. Softky and C. Koch, The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps, *J. Neurosci.*, **13** (1993), 334–350.

[49] P. Suksompong and T. Berger, Capacity analysis for integrate-and-fire neurons with descending action potential thresholds, *IEEE Trans. Inf. Theory*, **56** (2010), 838–851.

[50] P. J. Thomas and A. W. Eckford, Capacity of a simple intercellular signal transduction channel, preprint, `arXiv:1411.1650`.

[51] H. C. Tuckwell, *Introduction to Theoretical Neurobiology, Vol. 2*, Cambridge University Press, New York, 1988.

[52] S. Verdu, On channel capacity per unit cost, *IEEE Trans. Inf. Theory*, **36** (1990), 1019–1030.

[53] S. Verdu and T. S. Han, A general formula for channel capacity, *IEEE Trans. Inf. Theory*, **40** (1994), 1147–1157.

[54] A. D. Wyner, Capacity and error exponent for the direct detection photon channel – part I, *IEEE Trans. Inf. Theory*, **34** (1988), 1449–1461.

[55] A. D. Wyner, Capacity and error exponent for the direct detection photon channel – part II, *IEEE Trans. Inf. Theory*, **34** (1988), 1462–1471.

E-mail address: `kostal@biomed.cas.cz`

E-mail address: `shinomoto@scphys.kyoto-u.ac.jp`