

## Second Order Phase Transition in Neural Rate Coding: Binary Encoding is Optimal for Rapid Signal Transmission

Matthias Bethge,\* David Rotermund, and Klaus Pawelzik

*Institute of Theoretical Physics, University of Bremen, Otto-Hahn-Allee, D-28334 Bremen, Germany*

(Received 5 August 2002; published 27 February 2003)

Here, we derive optimal tuning functions for minimum mean square reconstruction from neural rate responses subjected to Poisson noise. The shape of these tuning functions strongly depends on the length  $T$  of the time window within which action potentials (spikes) are counted in order to estimate the underlying firing rate. A phase transition towards pure binary encoding occurs if the maximum mean spike count becomes smaller than approximately three. For a particular function class, we prove the existence of a second-order phase transition. The analytically derived critical decoding time window length is in precise agreement with numerical results. Our analysis reveals that binary rate encoding should dominate in the brain wherever time is the critical constraint.

DOI: 10.1103/PhysRevLett.90.088104

PACS numbers: 84.35.+i, 64.60.-i, 87.18.Sn

The striking feature of neurons in the central nervous system is the use of action potentials as a means of communication. Correspondingly, the coding properties of neurons are frequently characterized by tuning curves, which display the average firing rate of neurons as a function of the stimulus parameters of interest. This way of description is closely related to the idea of analog coding, which constitutes the basis for many neural network models. Apart from the success in stimulus reconstruction experiments, the idea of rate coding can also be justified by a basic biophysical property of neurons, namely, the temporal integration of the postsynaptic cells over the current pulses induced by the presynaptic spikes. Reliable inference from the observed number of spikes about the underlying firing rate of a neuronal response, however, requires a sufficiently long time interval, while integration times of neurons *in vivo* [1] as well as reaction times of humans or animals when performing classification tasks [2,3] are known to be rather short. Therefore, it is important to understand how neural rate coding is affected by the limited time window, which in reality is available for decoding.

In 1996, Softky pointed out that there is a trade-off between the higher information content of each analog “message” and the lower rate at which this message may be sent [4], so that the question arises how relevant the idea of analog coding actually is for neuronal processing in the brain. Although this is an important problem that even may be decidable experimentally, it did not receive much attention in neuroscience until today.

Here, we seek for optimal tuning functions that minimize the minimum average squared reconstruction error for a uniform source signal transmitted through a Poisson channel as a function of the maximum mean number of spikes. This tuning function approach appears to be particularly tailored to the issue of neural coding while, from a technical point of view, our work

contributes to a rate-distortion theory for the Poisson channel [5].

In formal terms, the issue is to optimally encode a real random variable  $x$  in the number of pulses emitted by a neuron within a certain time window. Thereby,  $x$  stands for the analog signal computed by a presynaptic neuron that shall be transmitted to subsequent neurons. The neuronal output actually read out by subsequent neurons, however, is given by the number of spikes  $k$  integrated within a time interval of length  $T$ . The statistical dependency between  $x$  and  $k$  then is specified by the assumption of Poisson noise [6]

$$p[k|\mu(x)] = \frac{[\mu(x)]^k}{k!} \exp\{-\mu(x)\}, \quad (1)$$

and the choice of the tuning function  $f(x)$ , which together with  $T$  determines the mean spike count  $\mu(x) = Tf(x)$ . An important additional constraint is the limited dynamic range of the neuronal firing rate, which can be included by the requirement of a bounded tuning function [ $f_{\min} \leq f(x) \leq f_{\max}$ ,  $\forall x$ ]. Since inhibition can reliably prevent a neuron from firing, we here consider the important case  $f_{\min} = 0$  only. Instead of specifying  $f_{\max}$ , we impose a bound directly on the mean spike count [i.e.,  $\mu(x) \leq \bar{\mu}$ ], because  $f_{\max}$  constitutes a meaningful constraint only in conjunction with a fixed time window length  $T$ .

As an objective function measuring the distortion, we consider the *minimum mean squared error* (MMSE) with respect to the uniform distribution  $\rho(x) = 1$  for  $x \in [0, 1]$ :

$$\begin{aligned} \chi^2[\mu(x)] &= E[x^2] - E[\hat{x}^2] \\ &= \frac{1}{3} - \sum_{k=0}^{\infty} \frac{\int_0^1 x p[k|\mu(x)] dx}{\int_0^1 p[k|\mu(x)] dx}, \end{aligned} \quad (2)$$

where  $\hat{x}(k) = E[x|k]$  denotes the *mean square estimator*, which is the conditional expectation (see, e.g., [8]). The exact shape of the optimal tuning function clearly depends on the distribution of  $x$ . Since natural signals are confined to a finite range, we choose the uniform distribution for the sake of symmetry and simplicity. The existence of the phase transition, however, appears to be independent of the shape of  $\rho(x)$ . In particular, we observed the same qualitative dependence in the case of other unimodal distributions [we checked for  $\rho(x) = (\nu + 1)2^\nu(0.5 - |0.5 - x|)^\nu$  with  $\nu = 0.2, 0.5, 1, 2, 4$ , and 10 exhibiting a phase transition at  $\bar{\mu}^c = 2.9, 2.9, 2.8, 2.7, 2.6$ , and 2.5, respectively].

Since in reality  $f_{\max}$  is bounded and fixed,  $\bar{\mu} = f_{\max}T$  is directly related to the rate  $1/T$  at which independent signals can be transmitted. Hence, in our study the trade-off between the higher information content of each analog message and the lower rate at which this message may be sent corresponds to the larger amount of time  $T$  that is necessary to achieve a lower distortion  $\chi^2$  by increasing the range of analog signaling.

As derived in [9] on the basis of Fisher information, the optimal tuning function for a single neuron in the asymptotic limit  $\bar{\mu} \rightarrow \infty$  has a parabolic shape [i.e.,  $f^{\text{asympt}}(x) = f_{\max}x^2$ ]. For any finite  $\bar{\mu}$ , however, this tuning function is not necessarily optimal and, in the limit  $\bar{\mu} \rightarrow 0$ , it is straightforward to show that the optimal tuning curve is a step function

$$f^{\text{step}}(x|\vartheta) = f_{\max} \Theta(x - \vartheta), \quad (3)$$

where  $\Theta(z)$  denotes the Heaviside function that equals one, if  $z > 0$  and zero if  $z < 0$ . The optimal threshold  $\vartheta(\bar{\mu})$  of the step tuning curve depends on  $\bar{\mu}$  and can be determined analytically:

$$\vartheta(\bar{\mu}) = 1 - \frac{3 - \sqrt{8e^{-\bar{\mu}} + 1}}{4(1 - e^{-\bar{\mu}})}, \quad (4)$$

as well as the corresponding MMSE [10]:

$$\chi^2[f^{\text{step}}] = \frac{1}{12} \left( 1 - \frac{3\vartheta^2(\bar{\mu})}{\{[1 - \vartheta(\bar{\mu})](1 - e^{-\bar{\mu}})\}^{-1} - 1} \right). \quad (5)$$

The binary shape for small  $\bar{\mu}$  and the continuous parabolic shape for large  $\bar{\mu}$  implies that there has to be a transition from discrete to analog encoding when  $\bar{\mu}$  is increased. Unfortunately, it is not possible to determine the optimal tuning function within the entire set of all bounded functions  $\mathcal{B} := \{f|f:[0, 1] \rightarrow [0, f_{\max}]\}$  using the calculus of variations. Instead we chose a certain parametrized function space  $S \subset \mathcal{B}$  in advance that is feasible for the optimization.

In this paper, we consider two function classes  $S_1, S_2$ , which both contain the binary tuning function as well as the asymptotic optimal parabolic function as special cases. Furthermore,  $S_1$  is a proper subset of  $S_2$ . By nu-

merical optimization within  $S_2$  for various  $\bar{\mu}$ , we found a clear phase transition from binary to analog encoding at a critical  $\bar{\mu}^c$  with  $2.9 < \bar{\mu}^c < 3.0$  (Fig. 1, upper panel). Although the critical value depends on the function space within which the optimization is performed, we did not find any tuning function with an error smaller than the MMSE of the step function for  $\bar{\mu} < 2.9$  (we investigated various function spaces, for which the results will be published elsewhere [10]).

Our interest in  $S_1$  results from the fact that we can analyze the phase transition in this subset analytically, while  $S_2$  is a quite large function space that is likely to sufficiently approximate all relevant tuning functions. Altogether  $S_2$  has six free parameters  $a \leq b \leq c \in [0, 1]$ ,  $f_{\text{mid}} \in (0, f_{\max})$ ,  $\alpha, \beta \in [0, \infty)$  and the parametrization of the tuning functions is given by

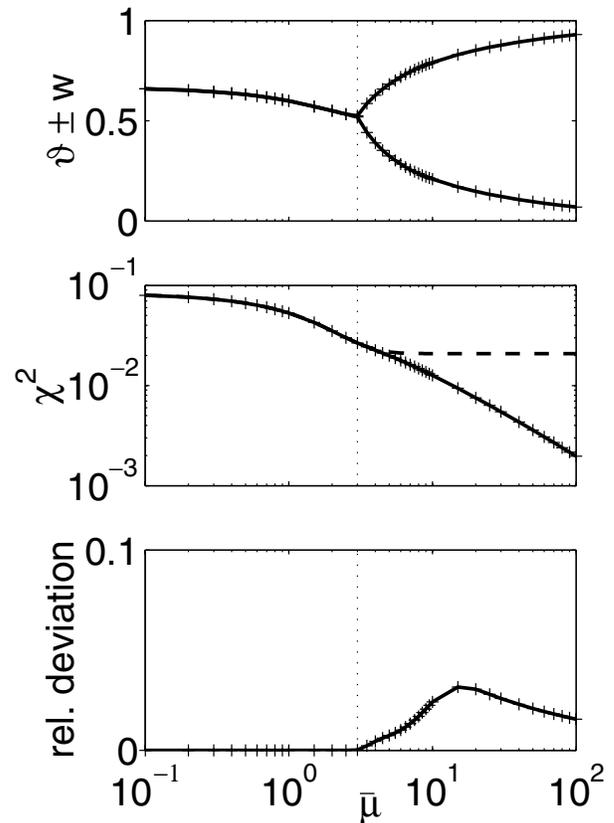


FIG. 1. The upper panel shows a bifurcation plot for  $\vartheta(\bar{\mu}) - w$  and  $\vartheta(\bar{\mu}) + w$  of the optimal tuning function in  $S_1$  illustrating the phase transition from binary to continuous encoding. The dotted line separates the regions before and after the phase transition in all three panels. Left from this line (i.e., for  $\bar{\mu} < \bar{\mu}^c$ ) the step function given by Eqs. (3) and (4) is optimal. The middle panel shows the MMSE of this step function (dashed line) and of the optimal tuning function in  $S_2$  (solid line), which becomes smaller than the first one after the phase transition. The relative deviation between the minimal errors of  $S_1$  and  $S_2$  [i.e.,  $(\chi_{S_1}^2 - \chi_{S_2}^2)/\chi_{S_2}^2$ ] is displayed in the lower panel.

$$f^{S_2}(x|a, b, c, f_{\text{mid}}, \alpha, \beta) = \begin{cases} 0 & 0 < x < a \\ f_{\text{mid}} \left(\frac{x-a}{b-a}\right)^\alpha & a < x < b \\ f_{\text{mid}} + (f_{\text{max}} - f_{\text{mid}}) \left(\frac{x-b}{c-b}\right)^\beta & b < x < c \\ f_{\text{max}} & c < x < 1 \end{cases} \quad (6)$$

Numerical optimization leads to the minimal MMSE as a function of  $\bar{\mu}$  as displayed in Fig. 1 (middle panel). The parametrization of the tuning functions in  $S_1$  is given by

$$f^{S_1}(x|w, \gamma) = \begin{cases} 0 & 0 < x < \vartheta(\bar{\mu}) - w \\ f_{\text{max}} \left(\frac{x - \vartheta(\bar{\mu}) + w}{2w}\right)^\gamma & \vartheta(\bar{\mu}) - w < x < \vartheta(\bar{\mu}) + w, \\ f_{\text{max}} & \vartheta(\bar{\mu}) + w < x < 1 \end{cases} \quad (7)$$

with  $w \in [0, 1]$  and  $\gamma \in [0, \infty)$ . The integrals entering Eq. (2) for the MMSE in case of the tuning function  $f^{S_1}$  read

$$\int_0^1 xp(k|x)dx = \frac{1}{k!} \left\{ \frac{[\vartheta(\bar{\mu}) - w]^2}{2} \delta_{0,k} + \frac{4w^2 \Gamma_{f_{\text{max}}}(k + \frac{2}{\gamma})}{\gamma (\sqrt[\gamma]{f_{\text{max}}})^2} \right. \\ \left. + \frac{2w[\vartheta(\bar{\mu}) - w] \Gamma_{f_{\text{max}}}(k + \frac{1}{\gamma})}{\gamma \sqrt[\gamma]{f_{\text{max}}}} \right. \\ \left. + \frac{1 - [\vartheta(\bar{\mu}) + w]^2}{2} f_{\text{max}}^k e^{-f_{\text{max}}} \right\}, \quad (8)$$

$$\int_0^1 p(k|x)dx = \frac{1}{k!} \left\{ [\vartheta(\bar{\mu}) - w] \delta_{0,k} + \frac{2w \Gamma_{f_{\text{max}}}(k + \frac{1}{\gamma})}{\gamma \sqrt[\gamma]{f_{\text{max}}}} \right. \\ \left. + [1 - \vartheta(\bar{\mu}) - w] f_{\text{max}}^k e^{-f_{\text{max}}} \right\}, \quad (9)$$

where  $\Gamma_v(z) = \int_0^v s^{z-1} e^{-s} ds$  denotes the truncated Gamma function. The minimal MMSE for these tuning functions is only slightly worse than that for  $S_2$ . The relative difference between both is plotted in Fig. 1 (lower panel) showing a maximum deviation of 3.2%.

In particular, the relative deviation is extremely small around the phase transition. This comparison suggests that a restriction to  $S_1$ , which is a necessary simplification for the following analytical investigation, does not change the qualitative results.

The phase transition from binary to analog encoding corresponds to a structural change of the objective function  $\chi^2(w, \gamma)$ . In particular, the optimality of binary encoding for  $\bar{\mu} < \bar{\mu}^c$  implies that  $\chi^2(w, \gamma)$  has a minimum at  $w = 0$ . The existence of a phase transition implies that with increasing  $\bar{\mu}$  this minimum changes into a local maximum at a certain critical point  $\bar{\mu} = \bar{\mu}^c$ . Therefore, the critical point can be determined by a local expansion of  $\chi^2(w, \gamma, \bar{\mu}) - \chi^2(0, \gamma, \bar{\mu}) = \sum_{k=1}^{\infty} g_k(\gamma, \bar{\mu}) \frac{w^k}{k!}$  around  $w = 0$ , because the sign of its leading coefficient  $A_\gamma(\bar{\mu})$  (i.e., the coefficient  $g_k$  with minimal  $k$  that does not vanish identically) determines whether  $\chi^2(w, \gamma, \bar{\mu})$  has a local minimum or maximum at  $w = 0$ . Accordingly, the critical point is given as the solution of  $A_\gamma(\bar{\mu}) = 0$ .

With quite a bit of effort, one can prove that the first derivative of  $\chi^2(w, \gamma, \bar{\mu})$  vanishes for all  $\bar{\mu}$ . The second derivative, however, is a decreasing function of  $\bar{\mu}$  and, hence, constitutes the wanted leading coefficient,

$$A_\gamma(\bar{\mu}) = \frac{1}{4(e^{\bar{\mu}} - 1)^2} \left\{ 8 - 7e^{\bar{\mu}} + 16e^{2\bar{\mu}} + e^{3\bar{\mu}} - \sqrt{1 + 8e^{-\bar{\mu}}} \{ 2 + e^{\bar{\mu}}[-3 + e^{\bar{\mu}}(6 + e^{\bar{\mu}})] \} \right. \\ \left. + \{ 16e^{\bar{\mu}} - 48e^{2\bar{\mu}} - 4e^{3\bar{\mu}} + \sqrt{1 + 8e^{-\bar{\mu}}} [4e^{\bar{\mu}} - 8(4 + e^{\bar{\mu}})] \} \frac{\bar{\mu}^{-(1/\gamma)}}{\gamma} \Gamma_{\bar{\mu}} \left( \frac{1}{\gamma} \right) \right. \\ \left. + [8e^{2\bar{\mu}} + 2(5 - 3\sqrt{1 + 8e^{-\bar{\mu}}})e^{3\bar{\mu}}] \frac{\bar{\mu}^{-(2/\gamma)}}{\gamma^2} \Gamma_{\bar{\mu}}^2 \left( \frac{1}{\gamma} \right) - 16e^{\bar{\mu}}(e^{\bar{\mu}} - 1)(\sqrt{1 + 8e^{-\bar{\mu}}} - 3) \right. \\ \left. \times \frac{\bar{\mu}^{-(2/\gamma)}}{\gamma} \Gamma_{\bar{\mu}} \left( \frac{2}{\gamma} \right) + 2e^{2\bar{\mu}}(e^{\bar{\mu}} - 1)(\sqrt{1 + 8e^{-\bar{\mu}}} - 3) \frac{\bar{\mu}^{-(2/\gamma)}}{\gamma^2} \right. \\ \left. \times \int_0^{\bar{\mu}} e^{-s} s^{(1-\gamma)/\gamma} \left( 1 - \frac{s}{\bar{\mu}} \right)^{-(1/\gamma)} \Gamma_{-s\bar{\mu}} \left( \frac{1}{\gamma} \right) ds \right\}. \quad (10)$$

Obviously, it is not possible to write the zeros of  $A_\gamma(\bar{\mu})$  in a closed form. The numerical evaluation of the critical point  $\bar{\mu}^c(\gamma)$  as a function of  $\gamma$  is displayed in Fig. 2. Note that we have treated  $\gamma$  as a fixed parameter, which means that we determine the critical point of the phase transition in all subsets  $S_1(\gamma)$  of  $S_1$  that correspond to a fixed  $\gamma$ . It is straightforward to show that the critical point  $\bar{\mu}^c$  with respect to the entire class  $S_1$  is given by the minimum of  $\bar{\mu}^c(\gamma)$ .

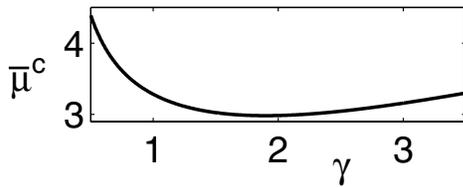


FIG. 2. The critical maximum mean spike count  $\bar{\mu}^c$  is shown as a function of  $\gamma$  (numerical evaluation at  $\gamma \in \{0.5, 0.505, 0.51, \dots, 3.5\}$ ). The minimum  $\bar{\mu}^c = 2.98291 \pm 10^{-7}$  at  $\gamma = 1.9$  determines the phase transition in  $S_1$ .

We determined this value up to a precision of  $\pm 0.0001$  to be  $\bar{\mu}^c = 2.9857$ . The shape of the optimal tuning function before and after the phase transition is displayed in Fig. 3. Since the structural change of the objective function is governed by a change of the sign of the leading coefficient, we thus have found a second-order phase transition.

Our study reveals that optimal encoding with respect to the minimum mean squared error is binary for maximum mean spike counts smaller than approximately three. Within the function class  $S_1$ , we determined a second-order phase transition from binary to continuous encoding analytically. With respect to mutual information, the advantage of binary encoding holds even up to a maximum mean spike count of about 3.5 (results not shown). Since the firing rates of cortical neurons are low relative to their short integration times, these results suggest that neural encoding in cortex should in fact be binary. In contrast, most experimental studies do not report on binary tuning functions, but instead show smooth tuning curves only. However, the shape of a tuning function always depends on the stimulus set used. Only recently, experimental studies under natural stimulus conditions provided evidence for the idea that neuronal encoding is essentially binary [11]. Particularly striking is this observation for the H1 neuron of the fly [12], for which the coding properties are probably better understood than for most other neurons [13]. In [12], it is remarkable that the neuronal response is the more binary the less noisy the stimulus conditions are (the noise level is determined by the different light conditions at midday, half an hour before, and half an hour after sunset), suggesting that the single neuron in fact does not encode more than two different states of the stimulus at a time.

Additional support for the potential relevance of a binary neural code comes from intracellular recordings *in vivo* revealing that the subthreshold membrane potential of many cortical cells switches between up and down states [14] depending on the stimulus. Furthermore, the dynamics of bursting cells plays an important role for

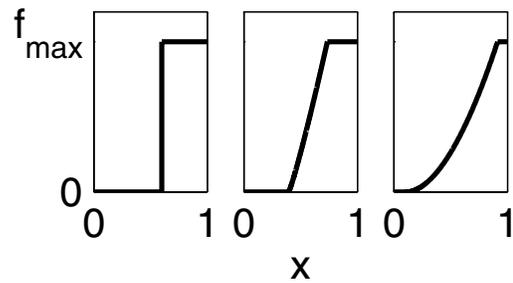


FIG. 3. Examples of the optimal tuning curve within  $S_1$  for  $\bar{\mu} < \bar{\mu}^c$  (left),  $\bar{\mu} = 5$  (middle), and  $\bar{\mu} = 50$  (right).

neuronal signal transmission [15] and may also be seen as evidence for binary rate coding. In light of these experimental facts, we conclude from our results that the idea of binary tuning constitutes an important hypothesis for neural coding that should be tested experimentally.

This work was supported by the DFG (SFB 517, Neurocognition).

---

\*Electronic address: mbethge@physik.uni-bremen.de  
URL: <http://neuro.physik.uni-bremen.de/~mbethge>

- [1] W. Softky and C. Koch, *J. Neurosci.* **13**, 334 (1993).
- [2] C. Keysers, D. Xiao, P. Foldiak, and D. Perrett, *J. Cog. Neurosci.* **13**, 90 (2001).
- [3] S. Thorpe, D. Fize, and Marlot, *Nature (London)* **381**, 520 (1996).
- [4] W. Softky, *Neural Networks* **9**, 15 (1996).
- [5] K. Rose, *IEEE Trans. Inf. Theory* **40**, 1939 (1994).
- [6] Poisson noise is the standard choice for neural rate coding [7]. More precisely,  $p[k|\mu(x)]$  denotes the conditional probability mass function of observing  $k$  given its mean value  $\mu$ , which itself is a function of  $x$ .
- [7] P. Dayan and L. Abbott, *Theoretical Neuroscience* (MIT Press, Cambridge, 2001).
- [8] E. Lehmann and G. Casella, *Theory of Point Estimation* (Springer, New York, 1999).
- [9] M. Bethge, D. Rotermund, and K. Pawelzik, *Neural Comput.* **14**, 2317 (2002).
- [10] M. Bethge, D. Rotermund, and K. Pawelzik (to be published).
- [11] P. Reinagel, *Curr. Opin. Neurobiol.* **11**, 437 (2001).
- [12] G. Lewen, W. Bialek, and R. de Ruyter van Steveninck, *Network: Comput. Neural Syst.* **12**, 317 (2001).
- [13] S. Strong, R. Koberle, R. de Ruyter van Steveninck, and W. Bialek, *Phys. Rev. Lett.* **80**, 197 (1998).
- [14] J. Anderson, I. Lampl, I. Reichova, M. Carandini, and D. Ferster, *Nature Neurosci.* **3**, 617 (2000).
- [15] J. Lisman, *Trends Neurosci.* **20**, 38 (1997).