Optimal tuning curves for neurons spiking as a Poisson process

Nicolas Brunel and Jean-Pierre Nadal

Laboratoire de Physique Statistique de l'E.N.S.,*
Ecole Normale Supérieure
24, rue Lhomond, F-75231 Paris Cedex 05, France
Tel.(33) 1.44.32.34.73 - 1.44.32.32.75 — Fax (33) 1.44.32.34.33
E-mail brunel@lps.ens.fr, nadal@lps.ens.fr

Abstract. We calculate the information capacity of a neuron emitting as a Poisson process in response to a static stimulus, and the stimulus distribution required to reach the capacity, in the case of a constraint on the average frequency. These optimal stimulus distributions (i.e. the ones reaching the information capacity) are then reexpressed in terms of 'tuning curves' for neurons with a continuous response to a scalar stimulus.

ESANN'97, Bruges (Belgium), 16-18 April 1997

1. Introduction

Since the work of Shannon[1] many researchers have tried to apply results of information theory to the study of the nervous system. Early papers focused on the information capacity of a spiking neuron using radically opposite hypothesis: Mackay and McCulloch[2] estimated the information capacity of a neuron using the duration of interspike intervals to encode information, while Stein[3] derived the information capacity of a neuron using a frequency code, obtaining estimates varying by several orders of magnitude. Until recent years however, very few attempts have been made to quantify the information transmitted by real nerve cells, and to attempt an explicit comparison with bounds provided by information theory [4, 5, 6].

In this paper we would like to perform a first step in making a comparison between bounds given by information theory, and information transmission by cells coding very simple stimuli, i.e. stimuli which can be characterized by a single scalar. The nervous system is full of examples of cells which seem to encode scalar stimuli with a frequency code: cells selective to orientation in primary visual cortex, to direction of motion in MT cortex, to head orientation in the subiculum of a rat, etc... For these systems, neurophysiologists have

---

[*] Laboratoire associé au C.N.R.S. (U.R.A. 1306), à l'ENS, et aux Universités Paris VI et Paris VII.

often determined 'tuning curves' with a reasonable precision. Thus one is tempted to translate a given tuning curve in distribution of frequencies, to estimate the mutual information given by a spike train in a given time interval on the stimulus, and compare it with the information capacity, or on the other hand, determine the optimal 'tuning curve' obtained from the distribution of frequencies that attain the information capacity. This is the goal of this paper. We have chosen for this preliminary study the simple situation of a static stimulus in a given interval, and of a neuron emitting spikes as a Poisson process, with a stimulus-dependent frequency.

The organization of the paper is the following: we first derive the information capacity of a spiking neuron with a constraint on the average frequency. This enables to determine the stimulus distribution optimizing the information transmitted per spike. Then we express these results in terms of tuning curves, and compare in conclusion the optimal tuning curves with experimental ones.

2. Information capacity and optimal stimulus distributions

We consider a single neuron emitting spikes according to a Poisson process: at a given frequency ν, after a time t, the number of emitted spikes k obeys to the distribution

$$Q(k \text{ spikes emitted in } [0,t] \mid \nu) = \frac{(\nu t)^k \exp(-\nu t)}{k!} \tag{1}$$

We assume that the frequency ν is a deterministic function of a scalar S which encodes the stimulus. If the stimulus is drawn randomly from a distribution $\rho(S)$, the frequency distribution $\mathcal{P}(\nu)$ is given by $\mathcal{P}(\nu) = \int dS\rho(S)\ \delta(\ \nu - \nu(S)\ )$

Some constraints may be added on the frequency distribution: a minimal, non zero, frequency, $\nu \geq \nu_{min}$, in order to take into account spontaneous activity; and a maximal value, $\nu \leq \nu_{max}$, in order to take into account the refractory period. The information capacity of such a neuron has been obtained by Stein[3]. Another type of constraint might be required, that of the *average* activity μ of a neuron.

The average information carried by the number of spikes observed during $[0,t]$ about the stimulus is

$$I(t) = \int d\nu \sum_k \mathcal{P}(\nu)\ Q(k|\nu) \log_2 \left(\frac{Q(k|\nu)}{p(k)} \right) \tag{2}$$

where the probability of observing k spikes given ν, $Q(k|\nu)$, is given by the Poisson law (1), and $p(k)$ is the probability that k spikes are emitted in $[0,t]$ averaged over the distribution of frequencies:

$$p(k) = \int d\nu \mathcal{P}(\nu)\ Q(k|\nu). \tag{3}$$

Another quantity of interest is the mutual information *per spike*, $i(t) = I(t)/(t\mu)$.

We have studied analytically two limits, the short time limit and the long time limit, and numerically the intermediate time regime.

2.1. Long times

Stein[3] showed that the information capacity is reached at long times when $\sqrt{\nu}$ has a uniform distribution on $[\sqrt{\nu_{min}}, \sqrt{\nu_{max}}]$. The capacity is

$$C(t) \;=\; \frac{1}{2}\log_2 \frac{2t\nu_{min}}{\pi e} \;+\; \log_2(\sqrt{\frac{\nu_{max}}{\nu_{min}}} \;-\; 1). \tag{4}$$

Alternatively the optimization can be considered under some additional constraint, such as a given mean value μ for the frequency. In that case one gets that the mutual information is maximized for $\sqrt{\nu}$ having the largest entropy under the constraint on the *square* of $\sqrt{\nu}$), and $\nu_{min} \leq \nu$: $\sqrt{\nu}$ has a truncated Gaussian distribution, which gives

$$\mathcal{P}_{opt}(\nu) \;=\; \frac{1}{Z\sqrt{2\pi\mu_0\nu}} \;\exp -\frac{\nu}{2\mu_0} \tag{5}$$

where Z and μ are constants (one has $\mu_0 \to \mu$ and $Z \to 1$ in the limit $\nu_{min} \to 0$).

The optimal mutual information is in this case

$$I(t) = \log_2 \sqrt{\frac{tZ^2\mu_0}{e}} + \frac{\mu}{2\mu_0 \log_2}, \quad \lim_{\nu_{min}\to 0} I(t) = \log_2 \sqrt{\frac{t\mu}{e}} + \frac{1}{2\log_2} \tag{6}$$

2.2. Short times

We consider now the opposite limit, that is the case $t \ll 1/\nu_{max}$, where ν_{max} is the highest possible frequency. In that limit, the mutual information, Eq. (2), is, at first order in t,

$$I(t) = t\int d\nu \mathcal{P}(\nu)\nu \;\log_2 \frac{\nu}{\mu} \tag{7}$$

where μ is the mean frequency. This first-order approximation has also been used by Skaggs et al[7] to obtain estimates of the information transmitted by the rat hippocampus about its environment. It can be shown to be an upper bound on $I(t)$ for any t[4].

Let us first consider the optimization under a given mean value μ. It is easy to check that a discrete, binary, frequency distribution maximizes the mutual information at this order in t under the constraint of a given mean frequency. For this optimal distribution, there are two possible frequencies, $\nu_1 = \nu_{min}$ and $\nu_2 = \nu_{max}$, occuring with probabilities P_1 and $P_2 = 1 - P_1$, respectively, with $P_2 = (\mu - \nu_{min})/(\nu_{max} - \nu_{min})$.

The optimal mutual information *per spike* $i(t)$ is given, in the simplest case $\nu_{min} = 0$, by

$$i(t) = \frac{I(t)}{\mu t} = \log_2 \frac{\nu_{max}}{\mu}$$

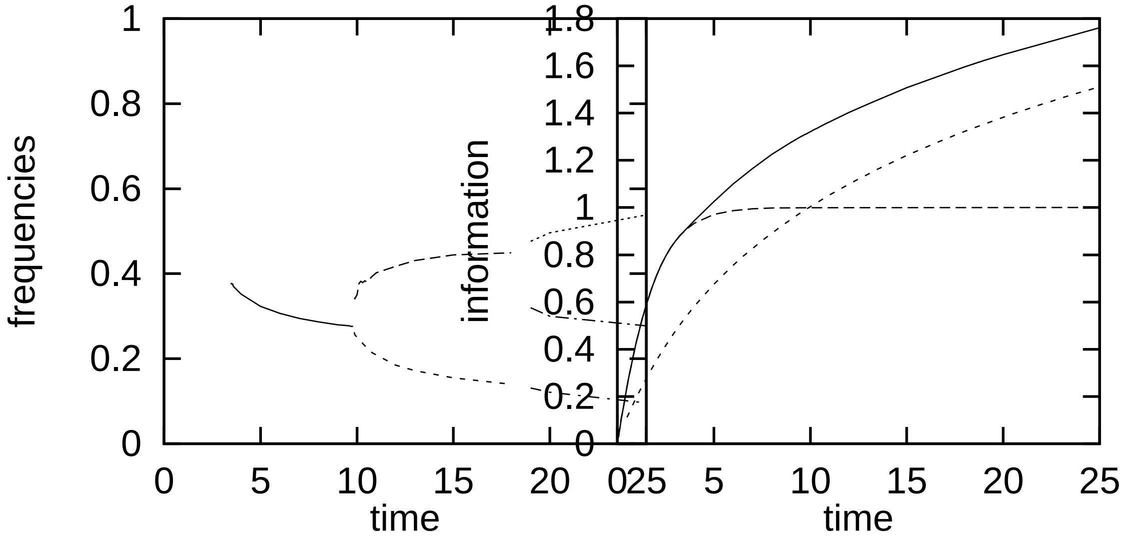

Figure 1: Left: Frequencies of the optimal distribution as a function of time. Right: Information capacity of the spiking neuron as a function of time (full curve); information transmitted in the case of a binary distribution (long-dashed curve); information transmitted in the case of the long-time optimal distribution (short-dashed curve)

which indicates that the optimal mutual information per spike is obtained in the limit $\mu \to \infty$, in which the mutual information goes to infinity.

Note that in this short time limit, for realistic values of ν_{max} and μ we typically obtain a mutual information per spike between 2 and 4 bits. This value is of the same order of magnitude as some estimates obtained with a time-varying stimulus[5].

2.3. Intermediate times

In the case of intermediate times we resorted to numerical analysis. In this section we consider the optimization with constraints on the minimal and maximal frequency only. Optimization with the additional constraint of an average frequency yields qualitatively similar results.

The results for $\nu_{min} = 0$, $\nu_{max} = 1$ are shown in Figs. 1. It shows transitions occurring at $t = 3.5, 9.8, 18.5\ldots$. At the first transition, $t \sim 3.5$, a peak at intermediate frequencies appears in the optimal distribution. At $t \sim 9.8$ a second transition occurs, at which a new peak appears. Increasing t further we find other transitions, with optimal distributions having an increasing number of peaks.

The information capacity in the interval $[0, t]$ is shown in Fig. 1. We show for comparison the information transmitted by a neuron with a binary distribution of frequencies, which saturates at 1 bit, and the information transmitted by a neuron with the optimal distribution at long times. It shows that the convergence to the optimal distribution at long times is rather slow. When t is small, the spiking neuron with the optimal distribution for long times tansmits about half of the information transmitted with the optimal binary distribution.

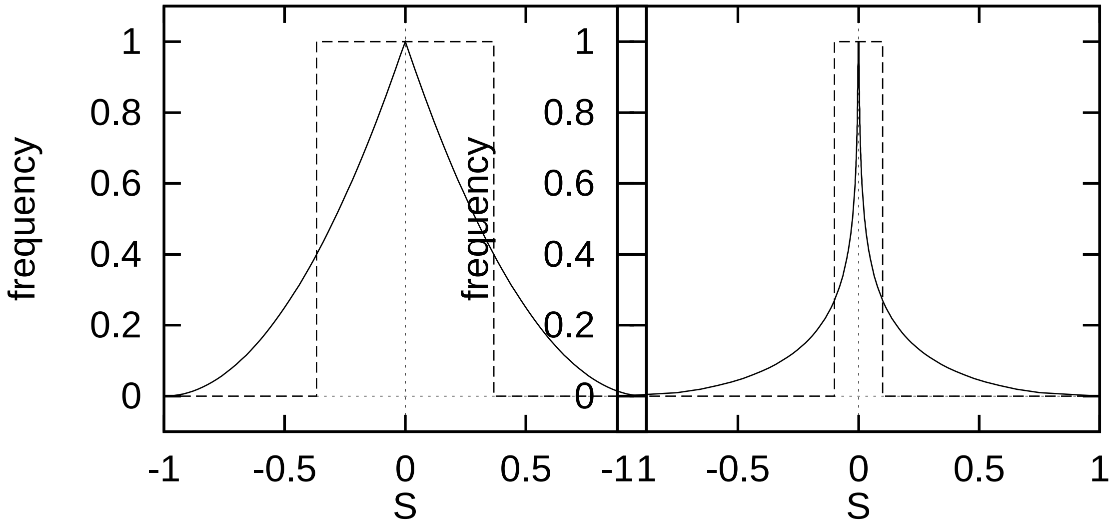

Figure 2: Left: Optimal 'tuning curves': long times (full curve), short times (dashed curve). Right: Optimal 'tuning curves' with constraint $\mu = 0.1$: long times (full curve), short times (dashed curve).

3. Optimal tuning curves

We now consider the case of a single neuron responding to a scalar stimulus, and make the connection between the results of the previous sections and 'tuning curves' more familiar to the neuroscience community. Suppose that the stimulus S is drawn randomly from a finite one dimensional domain $[-1, 1]$, with a uniform distribution. When stimulus S is present, the neuron responds with frequency $\nu = \phi(S)$. The curve $S \to \phi(S)$ defines the *tuning curve* of the neuron. The distribution of frequencies, expressed as a funtion of ϕ, is

$$\mathcal{P}(\nu) = \frac{1}{2}\int_{-1}^{1} dS\delta(\nu - \phi(S)) \tag{8}$$

We make the additional assumptions that the preferred stimulus is at $S = 0$ (i.e. $\phi(0) = \nu_{max}$); and that the tuning curve decreases monotonously from $\phi(0) = \nu_{max}$ to $\phi(1) = \nu_{min}$, and is symmetric around $S = 0$, i.e. $\phi(S) = \phi(-S)$.

With these assumptions we can obtain the optimal 'tuning curve' in the different cases studied in the previous section.

- **Long times, constraint on the maximal frequency**

$$\phi(S) = (\sqrt{\nu_{max}} - (\sqrt{\nu_{max}} - \sqrt{\nu_{min}})|S|)^2$$

- **Long times, constraint on the average frequency**

$$\phi(S) = \mu\left(H^{-1}\left(\frac{|S|}{2}\right)\right)^2 \quad \text{where} \quad H(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)dz$$

- **Short times**

$$\phi(S) = \begin{cases} \nu_{max} & \text{if } |S| < P_2 \\ \nu_{min} & \text{else} \end{cases} \tag{9}$$

These different tuning curves are shown in Fig. 2.

4. Conclusion

We have considered the information processing by neurons emitting spikes according to a Poisson process. We have shown that the stimulus distribution realizing the information capacity are very different at long or short times. Also, the mutual information is relatively much more sensitive to the distribution of frequencies for short times than for long times. Tuning curves that fit experimental data[10] seem closer to the optimal tuning curves that we have derived at short time. This is in agreement with the hypothesis that the nervous system must extract as much information as possible as fast as possible. There are however important differences between real tuning curves and the one predicted by our model. One can easily understand their origin. First, realistic tuning curves are expected to depend continuously on the stimulus strength. Second, such a continuous response function allows for the mutual information to be unbounded as time increases. Furthermore, in a wide range of t, tuning curves which fit experimental data lead to a mutual information close to the information capacity, at parity of minimal, maximal and average frequencies.

We have also obtained that the optimal information *per spike* is obtained in the short time limit for a binary sparse coding neuron. In order to exploit this information capacity neurons have to be able to respond in a very short time to incoming stimuli. Recent studies of recurrent networks of spontaneously active spiking neurons[8, 9] have shown that in presence of a balance between excitatory and inhibitory synaptic inputs, a neuron is able to respond very fast to small changes in the external inputs. This picture is consistent with neurophysiological and psychophysical experiments showing that processing of information in the visual cortex can occur at time scales of 10-20ms per layer.

References

[1] Shannon CE, *The Mathematical theory of communication*, (University of Illinois Press, Urbana, 1948)

[2] Mackay D and McCulloch W *Bull. Math. Biophys.* **14** 127

[3] Stein RB 1967 *Biophys. J.* **7** 797

[4] Bialek W and Zee A, *J. Stat. Phys.* **59** 103 (1989),

[5] Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D, *Science* **252** 1854 (1991),

[6] Bialek W, DeWeese M, Rieke F and Warland D, *Physica A* **200** 581 (1993)

[7] WE Skaggs, BL McNaughton, KM Gothard and EJ Markus, *NIPS*, 1993.

[8] Tsodyks M and Sejnowski T 1995 *Network* **6**

[9] Amit DJ and Brunel N 1996 *Cerebral Cortex*, in press

[10] Thorpe S, Fize D and Marlot C 1996 *Nature* **381** 520