

Coding of time-varying signals in spike trains of linear and half-wave rectifying neurons

Fabrizio Gabbiani†

Division of Biology, 139-74 California Institute of Technology, Pasadena, CA 91125, USA

Received 15 June 1995

Abstract. The encoding of time-varying stimuli in linear and half-wave rectifying neurons is studied. The information carried in single spike trains is assessed by reconstructing part of the stimulus using mean square estimation methods. For the class of models considered here, the mean square error in the reconstructions and estimates of the rate of information transmission are computed analytically. The optimal encoding of stimuli having statistical properties of natural images predicts a change in the temporal filtering characteristics with mean firing rate. This change relates to those observed experimentally at the early stages of visual processing. The transmission of information by model neurons is shown to be fundamentally limited to a maximum of 1.13 bit/spike and it is conjectured that nonlinear processing is necessary to explain higher rates which have been observed experimentally in certain preparations. In spite of the fact that single neurons might not transmit information efficiently, a substantial part of a time-varying stimulus can be recovered from single spike trains. In particular, our results demonstrate that a small number of ‘noisy’ neurons can carry precise temporal information in their spike trains.

1. Introduction

The neuronal coding and processing of sensory information is a subject which is presently being actively investigated [1, 22, 5, 30, 8]. Recently, methods of stochastic estimation theory [35, 24] have been applied to study the coding of random stimuli in the spike trains of various peripheral sensory neurons [5, 28]. By presenting repeatedly to an animal a stimulus drawn from a given probability distribution and recording the action potentials of a single cell, it is possible using stochastic estimation methods to reconstruct part of the stimulus from the recorded spike occurrence times, and hence to assess the information carried by the cell about the stimulus. The reconstruction algorithm and the more traditional Wiener kernel method [21] can be understood as complementary tools to investigate the information processing performed by a cell on its input signals. In both cases, one expresses the variable of interest as a Volterra series,

$$y = F_0 + F_1(z) + F_2(z, z) + \dots \quad (1.1)$$

where in the Wiener approach y is the mean firing rate of the cell (or the intracellular membrane voltage), while z is the stimulus, which consists of wide-band white noise (see figure 1(A), with $z = s$ and $y = O_M$). When applying the reconstruction algorithm, y is the stimulus which was presented to the cell and z is the spike train, from which one attempts to reconstruct the stimulus (see figure 1(B), with $z = x$ and $y = s_{\text{est}}$). The main goal

† E-mail address: gabbiani@klab.caltech.edu

of the Wiener approach is to compute the kernels F_0, F_1, F_2, \dots which provide a compact characterization of the processing performed by the neuron. The accuracy of the model obtained is assessed by computing the mean square error between the output (1.1) and the experimental output. It turns out that the kernels obtained from reconstructions are related in a complicated way to the encoding mechanisms of the cell (see section 5) and one is therefore mainly interested in the mean square error between the stimulus and estimated stimulus. As illustrated in figure 1(B), the mean square error in the reconstructions will depend on the preprocessing of the stimulus by the cell. If, for example, the stimulus contains frequencies higher than those encoded by the neuron, any information about them will be lost in the spike train and these will contribute significantly to the mean square error in the reconstructions (see section 2 and [11]). Hence, the choice of the bandwidth of the stimulus plays an important role and will significantly affect the calculated mean square error whereas it does not play any role when computing Wiener kernels (as long as the stimulus bandwidth is wide enough). Ideally, one would therefore like to pick a stimulus ensemble having a cut-off frequency matched to the frequencies that the cell encodes. While the bandwidth of a cell could be defined from a Wiener kernel model, we will show in sections 2 and 5 that an ‘effective bandwidth’ can also be determined from the signal-to-noise ratio of wide-band white noise reconstructions. It is then possible to measure the encoding capabilities of the cell by performing reconstructions of stimuli having a cut-off frequency matched to it.

An alternative way of characterizing reconstructions is to compute the mutual information rate between the experimental and estimated stimulus (1.1). The only quantity which can easily be computed, I_{LB} (see section 3 and appendix A; [32, 5, 28, 11]), is a lower bound on the rate of information transmission. It follows from the results derived in sections 3, 4 and 5 that comparing I_{LB} with a different information theoretic quantity, the ϵ -entropy (or rate-distortion function), I_ϵ allows one to draw some interesting conclusions on the ‘efficiency’ of single neurons to convey information on the presented stimulus.

The main subject of this work is to study the encoding of time-varying stimuli in the spike trains of simple model neurons by using the reconstruction method summarized above. We use a simplified integrate-and-fire neuron whose mean firing rate is proportional to a linearly filtered and half-wave rectified version of the stimulus $s(t)$. To implement the variability which is usually observed in response to several presentations of the same stimulus, the model possesses a random threshold (see figure 1(C)): the somatic current $q_s(t)$, which is obtained from $s(t)$ by linear filtering and half-wave rectification is integrated to yield the somatic membrane voltage (the spontaneous activity of the cell is set to zero). Once threshold is reached, a spike is generated, the voltage is reset to zero after a refractory period and a new threshold value is drawn from a fixed probability distribution. The simplification performed in section 4 consists of setting the refractory period δ equal to zero and choosing an exponentially distributed random threshold.

The choice of this model is motivated for two reasons. Firstly, there exists a number of neurons for which the approximation of linear filtering and half-wave rectification is expected to be meaningful. For example, the X-pathway in the cat (and monkey) visual system from retinal ganglion cells to simple cells in striate cortex has been shown to be satisfactorily described in this way (see, e.g., [29, 7, 10]), although deviations from linearity and half-wave rectification are expected [14, 25]. As shown in [11], this model also allows the investigation of some aspects of nonlinear processing. The second reason lies in mathematical simplicity: by choosing a vanishing refractory period and an exponentially distributed threshold it follows that the spikes of our neurons are generated independently of each other (i.e. are Poisson, see appendix B). Furthermore, since we consider only the

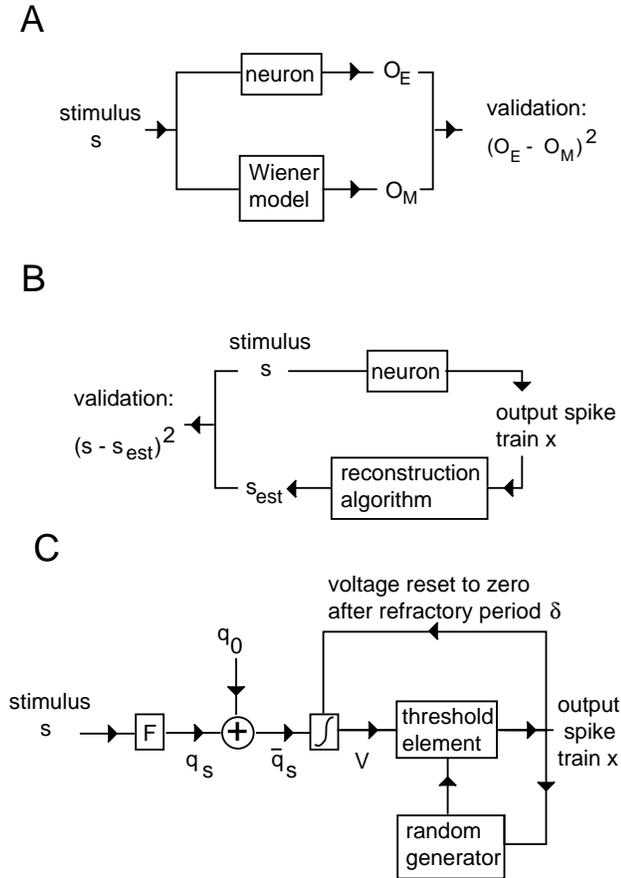


Figure 1. Schematic representation of the traditional Wiener kernel approach, the reconstruction method and the neuron model used in the subsequent sections. (A) In the Wiener approach, the experimental output is used to compute the kernels which constitute a model of the transformation performed by the neuron on the stimulus. The model validity is assessed by computing the mean square error between the experimental and predicted outputs. (B) In reconstructions, the output spike trains are used as an input to the algorithm which estimates the stimulus. Preprocessing of the stimulus by the neuron can strongly influence the outcome of the reconstructions. (C) The neuron model is an integrate-and-fire neuron with random threshold: the stimulus is filtered by F (we assume linear filtering and half-wave rectification) and the output q_s is added to q_0 (representing the background activity of the cell, in the following q_0 is set to zero). The somatic current \bar{q}_s is integrated to yield membrane voltage V and when the threshold is crossed a spike is fired, the voltage is reset to zero after a refractory period δ and the threshold is set to a new random value (the model used is a simplified version of this general model, see text for details).

encoding of Gaussian stimulus ensembles with zero mean, we can treat simultaneously the spike trains of two half-wave rectified cells, one encoding the positive and the other the negative variations of the stimulus (such as ON and OFF cells, for example) and, by using this trick, simplify the algebra. This allows us to compute exactly the cross-correlation between the spike trains and stimulus and the power spectrum of the spike trains (see section 4) and subsequently to investigate the encoding properties of these models analytically (see section 5). We show in particular that the reconstruction of a time-varying stimulus depends on the firing rate of the model neurons. While the mean square error

decreases with increasing mean firing rate, the information transmitted is fundamentally limited: it is not possible for the model neurons considered here to transmit more than $\pi/4 \log(2) \cong 1.13$ bit/spike of information about a time-varying white noise stimulus. These analytical results are only valid for the simplifications discussed above, but more general threshold distributions and stimulus ensembles, as well as a non-vanishing refractory period, could be investigated using computer simulations.

In section 6 two specific examples are considered: neuron models which low-pass filter their inputs and a simplified model of LGN cells. In the latter case we consider the encoding of Gaussian white ensembles as well as of stimuli having the power spectrum of natural visual scenes (as recently measured in [8]). This allows us to investigate the effect of noise on neuronal coding and to show that the optimal encoding of natural visual stimuli predicts a transition from low-pass to band-pass temporal filtering as a function of mean firing rate. Similar changes have been reported in the literature for retinal ganglion cells and LGN cells [31, 29].

The implications of these results and their relation to earlier work will be addressed in the concluding section of the paper. Most of the proofs will be relegated to appendix C.

2. Linear mean square estimation of a stimulus from a spike train

Let $s_0(t)$ be a random stimulus presented to the animal and $x_0(t)$ the spike train recorded from a cell. We assume that $s_0(t)$ and $x_0(t)$ are (real-valued) stochastic processes which are jointly (weakly-) stationary (that is, their joint one- and two-point correlation functions are time-translation invariant) and that $s_0(t)$ and $x_0(t)$ have finite variance, $\langle |s_0(t) - s_0|^2 \rangle = \langle |s_0(0) - s_0|^2 \rangle = \sigma^2 < \infty$, $\langle |x_0(t) - \lambda|^2 \rangle = \langle |x_0(0) - \lambda|^2 \rangle < \infty$, where $\lambda = \langle x_0(t) \rangle$ is the mean firing rate and $s_0 = \langle s_0(t) \rangle$ the mean value of the stimulus. In these equations, the brackets, $\langle \cdot \rangle$, denote average over the joint stimulus and spike train ensemble.

If $s(t) = s_0(t) - s_0$ and $x(t) = x_0(t) - \lambda$ are the stimulus and spike train with their mean value subtracted, we define the auto-correlations and cross-correlation of the stimulus and spike train as $R_{ss}(\tau) = \langle s(t)s(t+\tau) \rangle$, $R_{xx}(\tau) = \langle x(t)x(t+\tau) \rangle$ and $R_{sx}(\tau) = \langle s(t)x(t+\tau) \rangle$, respectively. A linear estimate $s_{\text{est}}(t)$ of the stimulus given to the spike train is obtained by setting

$$s_{\text{est}}(t) = \int_{-\infty}^{+\infty} dt_1 h(t_1) x(t - t_1)$$

for a square integrable function h . The filter h is to be chosen in such a way as to minimize the mean square error between the stimulus s and the estimate s_{est} ,

$$\epsilon^2 = \langle |s(t) - s_{\text{est}}(t)|^2 \rangle = \langle |s(0) - s_{\text{est}}(0)|^2 \rangle \quad (2.1)$$

where the last equality follows by stationarity. The orthogonality principle (see, e.g., [24], propositions V.C.2 and VII.C.1) implies that the optimal filter h satisfies the equation

$$R_{sx}(\tau) = (h \star R_{xx})(-\tau) \quad (2.2)$$

where the convolution $f \star g$ of two functions f and g is defined by

$$(f \star g)(t) = \int_{-\infty}^{+\infty} dt_1 f(t_1)g(t - t_1) = \int_{-\infty}^{+\infty} dt_1 f(t - t_1)g(t_1) = (g \star f)(t).$$

Equation (2.2) is solved by Fourier transformation. For a square integrable function h we define the Fourier transform of h through

$$\hat{h}(\omega) = \int_{-\infty}^{+\infty} d\tau h(\tau)e^{i\omega\tau} \quad h(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \hat{h}(\omega)e^{-i\omega\tau}$$

and for the auto-correlation and cross-correlation functions of the stimulus and spike train we set $S_{ss}(\omega) = \hat{R}_{ss}(\omega)$, $S_{xx}(\omega) = \hat{R}_{xx}(\omega)$ and $S_{sx}(\omega) = \hat{R}_{sx}(\omega)$. We restrict our attention to stimuli and spike train ensembles for which these functions are well defined and continuous. Furthermore, we assume the stimulus to be bandwidth limited, $S_{ss}(\omega) = 0$, if $|\omega| \geq \omega_c$, where $\omega_c = 2\pi f_c$ is the cut-off frequency and

$$0 < \epsilon < S_{ss}(\omega) < C \quad 0 < \epsilon < S_{xx}(\omega) < C \quad \text{if } \omega \in \Delta_c = (-\omega_c; +\omega_c)$$

for some positive constants ϵ and C .

From equation (2.2) we obtain the optimal filter h in Fourier space,

$$\hat{h}(\omega) = \begin{cases} \frac{S_{sx}(-\omega)}{S_{xx}(\omega)} & \text{if } \omega \in \Delta_c \\ 0 & \text{if } \omega \notin \Delta_c. \end{cases} \quad (2.3)$$

The solution (2.3) of the linear estimation problem is known as the non-causal Wiener filter (see, e.g., [24], section V.D.1). The filter h depends only on second-order statistical properties of the ensemble; hence different stimuli and spike train ensembles could lead to the same optimal linear decoding strategy.

Remark 2.1. (i) If the stimulus or the spike train sample functions contain deterministic components other than their mean value, these need to be subtracted as well in order for the preceding arguments to remain valid ([35], section 2.4).

(ii) We do not impose a causality constraint on the filter h although it could be implemented in several ways [11].

The difference between $s_{\text{est}}(t)$ and $s(t)$ is the ‘noise’ contaminating the reconstructions, $n(t) = s_{\text{est}}(t) - s(t)$. The noise has zero mean and its auto-correlation function is given by $R_{nn}(\tau) = R_{ss}(\tau) - (h \star R_{sx})(\tau)$ (using equation (2.2)). The power spectrum $S_{nn}(\omega) = \hat{R}_{nn}(\omega)$ is obtained by Fourier transformation,

$$S_{nn}(\omega) = S_{ss}(\omega) - \frac{|S_{sx}(\omega)|^2}{S_{xx}(\omega)} \quad (2.4)$$

using definition (2.3) of the filter h . If we define the signal-to-noise ratio as

$$SNR(\omega) = \frac{S_{ss}(\omega)}{S_{nn}(\omega)} \geq 1 \quad (2.5)$$

we can express the mean square error ϵ^2 in terms of the power spectrum of the stimulus and the signal-to-noise ratio:

$$\epsilon^2 = R_{nn}(0) = \frac{1}{2\pi} \int_{\Delta_c} d\omega \frac{S_{ss}(\omega)}{SNR(\omega)}. \quad (2.6)$$

The larger the signal-to-noise ratio, the smaller the mean square error. If, on the other hand, the signal-to-noise ratio is equal to 1 in some frequency band $\Delta \subset \Delta_c$, the entire power of the signal in this frequency band contributes to the mean square error. In the extreme case where the spike train is completely unrelated to the presented stimulus, $SNR(\omega) = 1$, $\forall \omega \in \Delta_c$, the mean square error coincides with the variance of the stimulus ensemble, $\epsilon^2 = \sigma^2$. Hence, the relative mean error, $\epsilon_r = \epsilon/\sigma$, represents an adequate measure of the accuracy of the reconstructions in the time domain, with $\epsilon_r = 1$ if the reconstruction of the signal is not better than chance and $\epsilon_r \rightarrow 0$ in the limit of perfect reconstructions.

Definition 2.1. Let the stimulus be wide-band white noise. If $\omega_{\max} \geq 0$ is the frequency at which the signal-to-noise ratio is maximal, $SNR(\omega_{\max}) \geq SNR(\omega)$ for $\omega \in \Delta_c$, then the *effective bandwidth of frequencies encoded by the cell* is defined as the interval $\Delta_e = (\omega_{\max} - \omega_{e1}; \omega_{\max} + \omega_{e2})$ around ω_{\max} for which

$$\frac{SNR(\omega) - 1}{SNR(\omega_{\max}) - 1} \geq r_{\min}.$$

The minimal value r_{\min} will be set to 0.05 in the following.

Remark 2.2. It will be shown in proposition 5.1 that, if $r_{\min} = 0.5$, definition 2.1 coincides with the usual definition of bandwidth for the model neurons studied in section 4.

3. Rate of information transmission

A lower bound on the rate, $I(s; s_{\text{est}})$, at which the reconstructions transmit information on the stimulus can be derived for $s(t)$ Gaussian. Since $s(t)$ was assumed to be bandwidth-limited, it is completely determined by the corresponding discrete time process $\tilde{s} = \{\tilde{s}_n\}_{n \in \mathbb{Z}}$, where

$$\tilde{s}_n = s\left(\frac{n\pi}{\omega_c}\right) \quad n \in \mathbb{Z}$$

according to Shannon's sampling theorem (see [15], theorem 2.6.2). Let $\tilde{s}_{\text{est}} = \{\tilde{s}_{\text{est}k}\}_{k \in \mathbb{Z}}$ and $\tilde{n} = \{\tilde{n}_k\}_{k \in \mathbb{Z}}$ denote the discrete time processes (sampled at the same frequency ω_c/π as \tilde{s}) corresponding to n and s_{est} and let $\tilde{n}^G = \{\tilde{n}_k^G\}_{k \in \mathbb{Z}}$ be the Gaussian process having the same covariance as $\tilde{n} = \{\tilde{n}_k\}_{k \in \mathbb{Z}}$. If we define $I_{\text{LB}} = \bar{h}(\tilde{s}) - \bar{h}(\tilde{n}^G)$, where $\bar{h}(\tilde{s})$ and $\bar{h}(\tilde{n}^G)$ are the entropy rate of \tilde{s} and \tilde{n}^G respectively, then $I(s; s_{\text{est}})$ is bounded below by

$$\begin{aligned} I(s; s_{\text{est}}) &= I(\tilde{s}; \tilde{s}_{\text{est}}) = \bar{h}(\tilde{s}) - \bar{h}(\tilde{n}|\tilde{s}_{\text{est}}) \\ &\geq \bar{h}(\tilde{s}) - \bar{h}(\tilde{n}) \geq I_{\text{LB}}. \end{aligned}$$

The last inequality follows because the entropy rate of a stationary process \tilde{n} is always smaller than the entropy rate of the corresponding stationary Gaussian process \tilde{n}^G (see [15], theorem 1.8.4).

From the entropy rate of a Gaussian process (see [15], theorem 2.4.1), we obtain

$$I_{\text{LB}} = \frac{1}{4\pi \log(2)} \int_{\Delta_c} d\omega \log \left(\frac{S_{ss}(\omega)}{S_{nn}(\omega)} \right) \quad (\text{in bit s}^{-1}). \quad (3.1)$$

Several equivalent formulae for this lower bound are known and are summarized in appendix A.

If the mean square error in the reconstructions is ϵ^2 , it will be of interest to compare I_{LB} to the absolute lower bound given by

$$I_\epsilon = \inf_{\rho} \{I(s; \rho) | \langle (s(t) - \rho(t))^2 \rangle \leq \epsilon^2\}.$$

I_ϵ is called the ϵ -entropy or rate-distortion function in information theory. By definition it follows that $I_{\epsilon_1} \leq I_{\epsilon_2} \Leftrightarrow \epsilon_1 \geq \epsilon_2$; this need not be the case for I_{LB} .

For a bandwidth-limited, stationary Gaussian process [16],

$$I_\epsilon = \frac{1}{4\pi \log(2)} \int_{\Delta_c} d\omega \log \max \left(\frac{S_{ss}(\omega)}{2\pi\theta^2}, 1 \right) \quad (\text{in bit s}^{-1}) \quad (3.2)$$

where the constant θ^2 is uniquely determined by

$$\int_{\Delta_c} d\omega \min \left(\frac{S_{ss}(\omega)}{2\pi}, \theta^2 \right) = \epsilon^2. \quad (3.3)$$

Equations (3.2) and (3.3) allow us to compute I_ϵ numerically, although a closed form expression will not be available in general.

A direct comparison of I_ϵ and I_{LB} is possible if the stimulus is white.

Proposition 3.1. Let the stimulus ensemble be bandwidth-limited Gaussian white noise and ϵ^2 be the mean square error in the reconstructions. Then

(i) the ϵ -entropy is given by the following formula:

$$I_\epsilon = \frac{-2f_c}{\log(2)} \log(\epsilon_r) \quad (\text{in bit s}^{-1})$$

(ii) $I_{LB} \geq I_\epsilon$.

Proof. See appendix C. □

In other words, the mutual information $I(s; s_{\text{est}})$ can be expected to be larger than the absolute lower bound determined by I_ϵ . I_{LB} and I_ϵ will be compared for several examples in section 6.

Finally, if λ is the mean firing rate of the cell, we define a lower bound on the rate of information transmitted per spike and the ϵ -entropy per spike according to $I_S = I_{LB}/\lambda$ and $I_{\epsilon S} = I_\epsilon/\lambda$.

4. Spike trains of linear and half-wave rectifying model neurons

In this section we study the spike trains generated by a simplified model whose mean firing rate is proportional to a linearly filtered and half-wave rectified version of the stimulus. It follows from earlier results [12] that this model is equivalent to an integrate-and-fire neuron with exponentially distributed random threshold and for which the refractory period has been set to zero. These results are summarized in appendix B. The assumption of a random threshold represents a convenient way to take into account the variability of spike trains to the presentation of identical stimuli which is usually observed experimentally. However, this assumption is not meant to imply that variability is solely due to the spike generating mechanism. In many cases, there is good evidence that this is not the case (see, for example, [20]).

Let $s(t)$ be Gaussian, as specified in the previous sections. We will encode a linearly filtered version of $s(t)$,

$$q_s(t) = \int_{-\infty}^{+\infty} dt_1 K(t - t_1)s(t_1) \quad \int_{-\infty}^{+\infty} dt_1 |K(t_1)|^2 < \infty \quad (4.1)$$

in the spike trains of two neurons, labelled ± 1 . The first neuron encodes the positive part of the signal, $q_s(t) > 0$ and the second neuron the negative part of the signal, $q_s(t) < 0$. We first formulate the encoding model for a finite time interval $(-T; T)$ and then compute the correlation functions in the limit $T \rightarrow \infty$.

Remark 4.1. (i) Causality requires $K(t) = 0$ for $t < 0$.

(ii) The Fourier transform $\hat{K}(\omega)$ of $K(t)$ will be assumed to be continuous and bounded, $|\hat{K}(\omega)| < C$, for $\omega \in \Delta_c$.

Let an event $e = (t_1, n_1; \dots, t_l, n_l)$ consist of a collection of time points $t_i \in (-T; T)$, $i = 1, \dots, l$, recording the occurrence of spikes, and for each time point t_i let n_i be a discrete variable, $n_i \in \{-1; 1\}$, specifying which of the two cells fired. The probability of an event e given the sample stimulus s is defined to be

$$\begin{aligned} P(e|s) &= Q_l(t_1, n_1; \dots, t_l, n_l|s) \\ &= C p(t_1, n_1|s) \cdots p(t_l, n_l|s) \quad Q_0(s) = C \end{aligned} \quad (4.2)$$

where the normalization constant C remains to be determined. In equation (4.2) we have made the assumption that each individual spike is generated independently of other spikes.

The probability density $p(t, \pm 1|s)$ is defined as

$$p(t, \pm 1|s) = \pm q_s(t)\theta(\pm q_s(t)) \quad (4.3)$$

with $q_s(t)$ given by equation (4.1) and where $\theta(\cdot)$ denotes the Heaviside function,

$$\theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0. \end{cases}$$

From definition (4.3) it follows that neuron 1 will have a non-vanishing probability density of firing at time t only if $q_s(t)$ is positive, whereas neuron -1 will have a non-vanishing probability density of firing only if $q_s(t)$ is negative.

It is easy to see that

$$p(t, 1|s) + p(t, -1|s) = |q_s(t)| \quad (4.4)$$

$$p(t, 1|s) - p(t, -1|s) = q_s(t). \quad (4.5)$$

The constant C in equation (4.2) is obtained from the normalization condition for the probability density functions Q_i ,

$$Q_0(s) + \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{n_1, \dots, n_i = \pm 1} \int \cdots \int_{-T}^T dt_1 \cdots dt_i Q_i(t_1, n_1; \dots, t_i, n_i|s) = 1. \quad (4.6)$$

By definition, each $Q_i(\cdot|s)$ is totally symmetric in its arguments. Plugging definitions (4.2) and (4.3) in (4.6) and using equation (4.4) to compute $\sum_{n_i = \pm 1} p(t_i, n_i|s)$, we obtain $C = e^{-\bar{N}}$, with $\bar{N} = \int_{-T}^T dt |q_s(t)|$.

The mean number of action potentials fired by neuron 1 in the interval $(-T; T)$ is given by

$$\begin{aligned} \bar{N}_1 &= e^{-\bar{N}} \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{n_1, \dots, n_i = \pm 1} \int \cdots \int_{-T}^T dt_1 \cdots dt_i \left(\sum_{j=1}^i \theta(n_j) \right) p(t_1, n_1|s) \cdots p(t_i, n_i|s) \\ &= e^{-\bar{N}} \sum_{i=1}^{\infty} \frac{\bar{N}^{i-1}}{(i-1)!} \left(\int_{-T}^T dt_1 q_s(t_1)\theta(q_s(t_1)) \right) = \int_{-T}^T dt_1 q_s(t_1)\theta(q_s(t_1)). \end{aligned}$$

Hence, the instantaneous firing rate of neuron 1 at time t is given by $\lambda_1 = q_s(t)\theta(q_s(t))$. Similarly, $\lambda_{-1} = -q_s(t)\theta(-q_s(t))$ is the instantaneous firing rate of neuron -1 and $\lambda_1 + \lambda_{-1} = |q_s(t)|$ the instantaneous firing rate of both cells under the driving stimulus s .

To each event $e = (t_1, n_1; \dots, t_l, n_l)$ we associate the sample $x(t, e)$,

$$x(t, e) = \sum_{i=1}^l n_i u(t - t_i) \quad x(t, e) = 0 \text{ if } l = 0 \quad (4.7)$$

of the stochastic process $x(t)$ which represents the spike trains of both cells. The function $u(t)$ describing a single spike is positive, square integrable and $\int_{-\infty}^{+\infty} dt u(t) = 1$. In subsequent sections we will consider the limit case $u(t) = \delta(t)$, where $\delta(t)$ is the Dirac delta function. Multiplication of each spike $u(t - t_i)$ by n_i in equation (4.7) allows us to distinguish which of the two cells fired at time t_i .

Proposition 4.1. Let $s(t)$ be a Gaussian stimulus ensemble, as specified in the preceding sections, and $x(t)$ the spike trains generated by two model neurons driven by the stimulus s . In the limit $T \rightarrow \infty$,

$$\langle s(t_a)x(t_b) \rangle = ((u \star K) \star R_{ss})(t_b - t_a) \quad (4.8)$$

$$\langle x(t_a)x(t_b) \rangle = \lambda_K(u \star u^r)(t_b - t_a) + [((u \star u^r) \star (K \star K^r)) \star R_{ss}](t_b - t_a) \quad (4.9)$$

where the averages on the left-hand side of equations (4.8), (4.9) are taken over the stimulus and spike train ensembles, we have defined $u^r(t) = u(-t)$, $K^r(t) = K(-t)$ and

$$\lambda_K = \langle |q_s(t)| \rangle = \frac{1}{\pi} \left(\int_{\Delta_c} d\omega |\hat{K}(\omega)|^2 S_{ss}(\omega) \right)^{1/2} \quad (4.10)$$

is the mean firing rate of both cells averaged over the stimulus ensemble.

Proof. See appendix C. □

5. Properties of linear reconstructions

The optimal non-causal Wiener filter (2.3) decoding the spike trains of our model neurons is obtained by Fourier transforming equations (4.8) and (4.9),

$$S_{sx}(\omega) = \hat{u}(\omega) \hat{K}(\omega) S_{ss}(\omega) \quad (5.1)$$

$$S_{xx}(\omega) = |\hat{u}(\omega)|^2 (\lambda_K + |\hat{K}(\omega)|^2 S_{ss}(\omega)) \quad (5.2)$$

so that choosing $u(t) = \delta(t)$, we obtain

$$\hat{h}(\omega) = \frac{\hat{K}(-\omega) S_{ss}(\omega)}{\lambda_K + |\hat{K}(\omega)|^2 S_{ss}(\omega)}. \quad (5.3)$$

The following result relates definition 2.1 with the usual notion of bandwidth.

Proposition 5.1. Let the stimulus ensemble be bandwidth-limited Gaussian white noise. The signal-to-noise ratio in the reconstructions satisfies the equation

$$\frac{SNR(\omega) - 1}{SNR(\omega_{\max}) - 1} = \frac{|\hat{K}(\omega)|^2}{|\hat{K}(\omega_{\max})|^2}. \quad (5.4)$$

Proof. This follows from equations (2.4), (2.5), (5.1) and (5.2). □

Remark 5.1. (i) It follows from equation (5.4) that definition 2.1 coincides with the usual definition of bandwidth for $r_{\min} = 0.5$ (i.e. Δ_c is the range of frequencies for which the attenuation in the gain is less than -3 dB).

(ii) Typically, single neurons never achieve the performance of electronic devices and the engineering cut-off criterion has to be relaxed when applied to such systems. With $r_{\min} = 0.05$ the cut-off criterion corresponds to an attenuation of -13 dB. The choice of a lower bound of 5% for the normalized signal-to-noise ratio in equation (5.4) is necessarily arbitrary, but proved to be appropriate in this theoretical work. In an experimental situation, one might want to choose a higher value depending on the accuracy of the numerical estimate for the signal-to-noise ratio.

The encoding of time-varying stimuli in our model neurons is characterized by the following three properties.

Property 5.2. Let the stimulus ensemble be Gaussian and bandwidth-limited. The mean square error, ϵ^2 , in the reconstructions is given by

$$\epsilon^2 = \frac{1}{2\pi} \int_{\Delta_c} d\omega \frac{\lambda_K S_{ss}(\omega)}{\lambda_K + |\hat{K}(\omega)|^2 S_{ss}(\omega)} \quad (5.5)$$

and is a monotone decreasing function of the mean firing rate λ_K with

$$\epsilon^2 \rightarrow \sigma^2 \quad (\lambda_K \rightarrow 0) \quad \text{and} \quad \epsilon^2 \rightarrow \mu^2 = \int_{\Delta} d\omega S_{ss}(\omega) \quad (\lambda_K \rightarrow \infty)$$

where $\Delta = \{\omega | \omega \in \Delta_c \text{ and } |\hat{K}(\omega)|^2 = 0\} \subset \Delta_c$.

Proof. See appendix C. \square

Remark 5.2. In the limit of low firing rates, the mean square error converges to the variance of the stimulus, which is the worst possible error (see section 2). As $\lambda_K \rightarrow \infty$ the model neurons are able to encode arbitrarily well the stimulus in the range of frequencies $\Delta_c - \Delta$ which are transmitted by the cell.

Property 5.3. Let the stimulus be Gaussian and bandwidth-limited.

(i) The lower bound, I_{LB} , on the rate of information transmission is given by

$$I_{LB} = \frac{1}{4\pi \log(2)} \int_{\Delta_c} d\omega \log \left(1 + \frac{1}{\lambda_K} |\hat{K}(\omega)|^2 S_{ss}(\omega) \right) \quad (\text{in bit s}^{-1}). \quad (5.6)$$

As a function of the mean firing rate λ_K , I_{LB} is a monotone increasing function satisfying

$$I_{LB} \rightarrow 0 \quad (\lambda_K \rightarrow 0) \quad \text{and} \quad I_{LB} \rightarrow \infty \quad (\lambda_K \rightarrow \infty).$$

(ii) The lower bound on the information transmitted per spike, I_S , is a monotone decreasing function of the mean firing rate λ_K with

$$I_S \rightarrow \frac{\pi}{4 \log(2)} \cong 1.13 \text{ bit/spike} \quad (\lambda_K \rightarrow 0) \quad \text{and} \quad I_S \rightarrow 0 \quad (\lambda_K \rightarrow \infty). \quad (5.7)$$

Proof. See appendix C. \square

Remark 5.3. Remarkably, the limit $\lambda_K \rightarrow 0$ is independent of the statistics of the stimulus ensemble $S_{ss}(\omega)$ and of the linear filter $K(t)$ preceding the spiking mechanism of our model neurons.

Property 5.4. Let the stimulus ensemble be bandwidth-limited Gaussian white noise and ϵ^2 be the mean square error in the reconstructions.

(i) As a function of the mean firing rate, λ_K , I_ϵ is monotone increasing with

$$I_\epsilon \rightarrow 0 \quad (\lambda_K \rightarrow 0) \quad I_\epsilon \rightarrow \frac{-\omega_c}{2\pi \log(2)} \log \left(\frac{\mu^2}{\sigma^2} \right) \quad (\lambda_K \rightarrow \infty).$$

(ii) The ϵ -entropy per spike, $I_{\epsilon S}$, is a monotone decreasing function of the mean firing rate λ_K with

$$I_{\epsilon S} \rightarrow \frac{\pi}{4 \log(2)} \cong 1.13 \text{ bit/spike} \quad (\lambda_K \rightarrow 0) \quad I_{\epsilon S} \rightarrow 0 \quad (\lambda_K \rightarrow \infty). \quad (5.8)$$

Proof. See appendix C. \square

Remark 5.4. (i) In contrast to I_{LB} , I_ϵ remains finite when some frequency band of the stimulus ensemble is not encoded.

(ii) For Gaussian white noise I_ϵ and I_{LB} converge to the same limit as $\lambda_K \rightarrow 0$; this result will be discussed in section 7.

If the second-order statistics of the temporal stimuli to be encoded is known, it is of interest to determine which filter K preceding the spiking mechanism will minimize the mean square error in the reconstructions. This amounts to minimizing the functional

$$\epsilon^2(K) = \frac{1}{2\pi} \int_{\Delta_c} d\omega \frac{\lambda_K S_{ss}(\omega)}{\lambda_K + |\hat{K}(\omega)|^2 S_{ss}(\omega)}. \quad (5.9)$$

As shown by property 5.2, this minimization problem is only meaningful under the constraint that the firing rate of the cell λ_K is held constant, $\lambda_K = \lambda_0$, with λ_0 fixed.

Proposition 5.5. Let the stimulus be bandwidth-limited and Gaussian with power spectrum $S_{ss}(\omega)$. The optimal filter gain $|\hat{K}(\omega)|^2$ which minimizes the functional (5.9) under the constraint of a fixed mean firing rate λ_0 is given by

$$|\hat{K}(\omega)|^2(\omega) = \begin{cases} \lambda_K \frac{\alpha^{-1/2} S_{ss}(\omega)^{1/2} - 1}{S_{ss}(\omega)} & \text{if } \omega \in \Delta_1 \\ 0 & \text{otherwise} \end{cases} \quad (5.10)$$

where the Lagrange multiplier α and the domain $\Delta_1 = \{\omega | \alpha^{-1/2} S_{ss}(\omega)^{1/2} > 1\} \subseteq \Delta_c$ satisfy the constraint equation

$$\frac{1}{\pi^2} \int_{\Delta_1} d\omega S_{ss}(\omega) |\hat{K}(\omega)|^2(\omega) = \lambda_0. \quad (5.11)$$

Proof. See appendix C. □

Remark 5.5. We do not consider the problem of finding a causal filter $K_c(t)$ whose gain $|\hat{K}_c(\omega)|$ approximates (5.10). This problem has been extensively treated in the literature (see, for example, [23], section 12-1; [17], sections 14-8 and 14-9). An example showing how a causal solution $K_c(t)$ is computed from $|\hat{K}_c(\omega)|$ is treated in [8].

6. Examples

The results obtained in section 5 will now be applied to two specific examples. In both cases, we first compute the effective bandwidth of frequencies encoded by the models and then the relative error in the reconstructions of stimuli having a cut-off frequency f_c matched to the effective bandwidth. We impose a minimal cut-off frequency of 15 Hz for the stimuli when the bandwidth of model neurons is smaller than 15 Hz. The introduction of a lower bound on the frequency content of the stimulus prevents its time fluctuations from becoming arbitrarily slow.

Example 6.1. The simplest biologically relevant assumption for K consists of low-pass filtering [11]. The effective bandwidth of an exponential low-pass filtering neuron,

$$K(t) = \begin{cases} \alpha e^{-t/\tau} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (6.1)$$

is given by

$$\frac{SNR(\omega) - 1}{SNR(0) - 1} = \frac{1}{1 + \omega^2 \tau^2}$$

($\omega_{e1} = \omega_{\max} = 0$) and with $f_e = \omega_{e2}/2\pi$ the criterion (5.4) corresponds to $f_e = \sqrt{19}/2\pi \tau$ Hz. Table 1 gives f_e for values of τ which are expected to lie in the physiological

range. If the stimulus is Gaussian, white with a cut-off frequency ω_c and variance $\sigma^2 = R_{ss}(0)$, the mean square error ϵ^2 can be calculated using elementary integral identities:

$$\epsilon^2 = \frac{\sigma^2}{\omega_c} \left(\omega_c - \frac{1}{\tau} \frac{\gamma}{\sqrt{1+\gamma}} \arctan \left(\frac{\tau \omega_c}{\sqrt{1+\gamma}} \right) \right) \quad \gamma = \frac{\alpha^2 \tau^2 \sigma^2}{\omega_c} \quad (6.2)$$

(see equations (4.10) and (5.5)). The lower bound I_{LB} is obtained similarly after a partial fraction expansion and integrating by parts:

$$I_{LB} = \frac{1}{2\pi \log(2)} \left[\omega_c \log \left(1 + \frac{\gamma}{1 + \omega_c^2 \tau^2} \right) + \frac{2}{\tau} \sqrt{1+\gamma} \arctan \left(\frac{\tau \omega_c}{\sqrt{1+\gamma}} \right) - \frac{2}{\tau} \arctan \tau \omega_c \right]. \quad (6.3)$$

Table 1. Effective bandwidth $\Delta_c = (0; f_e)$ of low-pass filtering model neurons as a function of the time constant of low-pass filtering, τ .

τ (ms)	f_e (Hz)
10	69
20	35
50	14
100	7
200	3.5

The optimal filter encoding the white noise stimulus is obtained from proposition 5.2. Setting $\Delta_1 = \Delta_c$ and solving the constraint (5.11) for $\alpha^{1/2}$ we obtain from (5.10)

$$|\hat{K}(\omega)|^2 = \frac{\pi \lambda_0^2}{2\sigma^2} \quad \forall \omega \in \Delta_c. \quad (6.4)$$

It follows from equation (6.4) that for the optimal filtering neuron $I_{LB} = I_e$, whereas for low-pass filtering neurons, $I_{LB} > I_e$ (see figure 2(B)).

Figure 2(A) plots the relative mean error at various firing rates for low-pass filtering neurons with time constants $\tau = 10$ and 200 ms (see table 1) as well as for the optimal filtering neuron in response to a Gaussian white stimulus (see equations (6.1) and (6.4)). The cut-off frequency f_c is equal to f_e (figure 2(A), top panel) or set to 15 Hz when $f_e < 15$ Hz (figure 2(B), bottom panel). It can be concluded that a single low-pass filtering neuron encodes at least 10–30% of a time-varying white noise signal in its optimal frequency band. Furthermore, in the first case ($\tau = 10$ ms) the mean relative error compares well with the absolute lower bound set by the optimal filtering model. For a time constant of 200 ms, the effective bandwidth is strongly reduced (see table 1) and the performance starts to differ significantly from the ideal one. Nevertheless, a low-pass filtering neuron with a time constant of 200 ms is still able to encode between 10 and 30% of a time-varying white stimulus having a bandwidth of 15 Hz. As shown in figure 2(B), $I_e/2$ varies between 0 and a maximum of 40 bit s⁻¹ (for $\tau = 10$ ms and a firing rate of 100 Hz). As expected, the optimal filtering neuron always transmits more information about the stimulus than the low-pass filter model neurons and, just as for the mean relative error, the difference between the ϵ -entropies of these two models increases with τ , in parallel with the decrease in effective bandwidth. In the worst case, $\tau = 200$ ms, $I_e/2$ ranges from a minimum of 3 bit s⁻¹ to a maximum of 10 bit s⁻¹.

It follows from equation (6.4) that in order to reconstruct a white stimulus as accurately as possible, the linear filter $K(t)$ should pass all frequencies $\omega \in \Delta_c$ equally well. This is

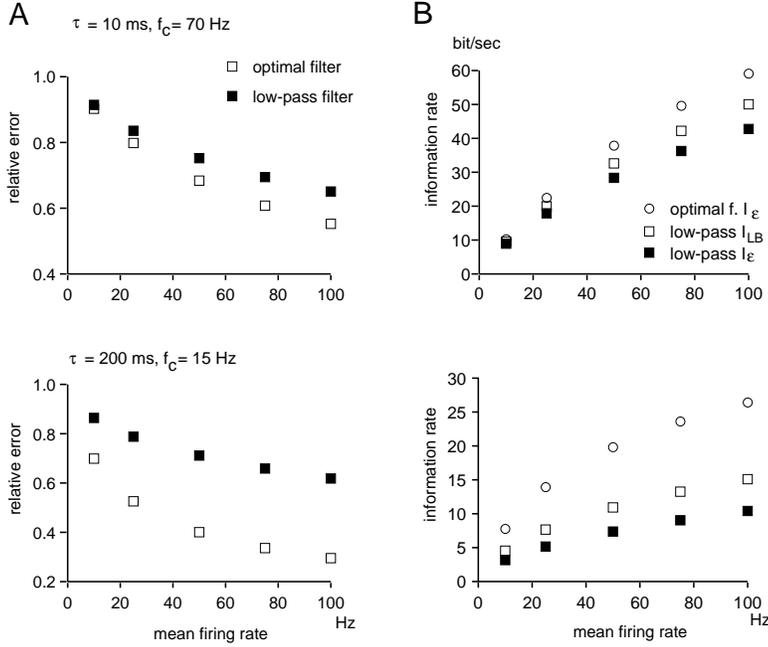


Figure 2. (A) Relative error (ϵ_r , filled squares) in the reconstructions of a white stimulus as a function of mean firing rate ($\lambda_K/2$) for two low-pass filtering model neurons. The cut-off frequency of the signal coincides with the effective bandwidth f_e for $\tau = 10$ ms and is set to 15 Hz for $\tau = 200$ ms. The empty squares report the relative mean error for the optimal encoding of the same white noise stimulus (see equation (6.4)). (B) Comparison of the lower bound for the rate of information transmission ($I_{LB}/2$, empty squares) and the ϵ -entropy ($I_\epsilon/2$, filled squares) for same low-pass filter models in response to a white stimulus. The circles report the lower bound for the rate of information transmission and the ϵ -entropy of the optimal filter model ($I_{LB} = I_\epsilon$) in response to the same stimulus.

consistent with the fact that the spike trains of our model neurons are Poisson processes; in this case the noise introduced by the spike generation mechanism is independent of frequency. If the power spectrum of the input signal is not white, the situation changes drastically, as is illustrated in the following.

Example 6.2. The white stimulus considered in example 6.1 is not likely to be typical of stimuli encountered in a natural environment. Although little information is presently available on the temporal statistics of natural visual or auditory stimuli, it has been recently reported [8] that the temporal power spectrum of natural time-varying images (in the limit of zero spatial frequency) obeys a quadratic power law decay $\sim 1/\omega^2$ similar to the decay in spatial power observed for static images [9]. This experimental result motivates our second example,

$$S_{ss}(\omega) = \frac{\mu}{1 + \tau^2 \omega^2} \quad (6.5)$$

where we have introduced a cut-off at low temporal frequencies to regularize the divergence $1/\omega^2 \rightarrow \infty$ as $\omega \rightarrow 0$. The time constant τ is chosen equal to 1500 ms, close to the value which approximates the power spectrum reported in [8] (see [8], figure 1, top). Note also that we do not take into account the noise term of [8], equation (13) since an ‘effective

noise' is already present in the spiking mechanism of our model neurons, see the discussion in section 7.

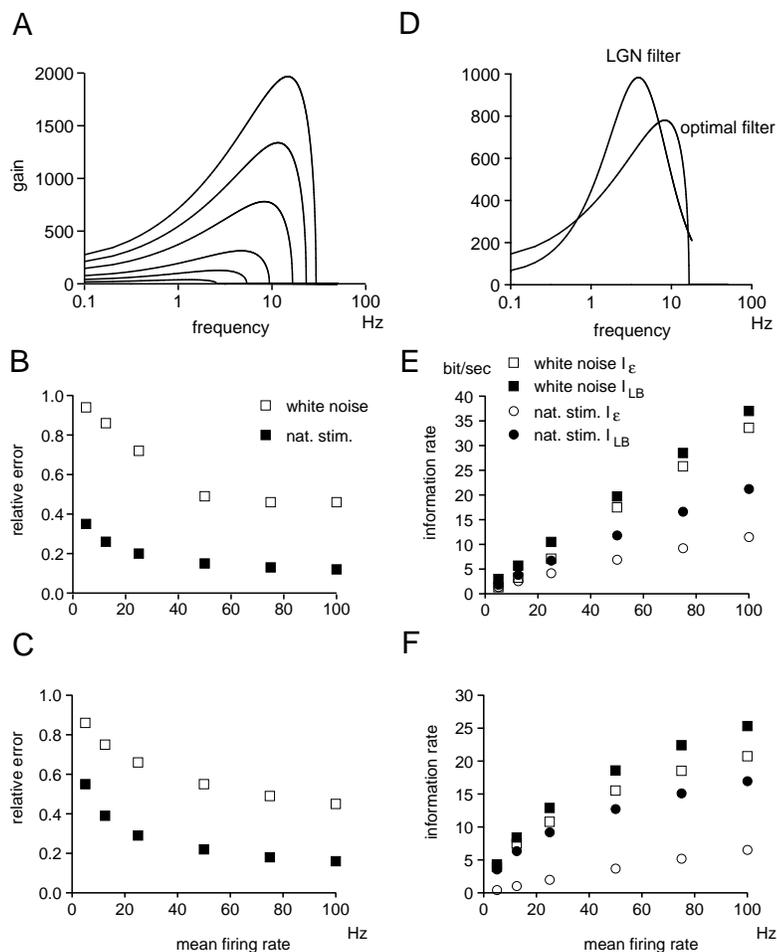


Figure 3. (A) Gain of the optimal filter model as a function of temporal frequency for a stimulus ensemble whose spectral power density decays quadratically with temporal frequency. The filters are shown for firing rates of $\lambda_K/2 = 5, 12.5, 25, 50, 75$ and 100 Hz (from the smallest to the largest gain), respectively (note the logarithmic temporal frequency scale; the units for the gain of the filter are arbitrary and correspond to $\mu = 1$, see main text). (B) Relative error (ϵ_r) in the reconstructions of a white (empty squares) and a natural stimulus (filled squares) for the optimal encoding filter model; the bandwidth of the input signal coincides with the effective bandwidth of the cell (which depends on the mean firing rate, see table 2; $f_c = 15$ Hz if $f_c < 15$ Hz). (C) Mean relative error for the gain of an LGN cell as measured experimentally [29] and fitted in [8]. The bandwidth of the stimulus is equal to the effective bandwidth of the cell (18 Hz). (D) Comparison between the LGN gain and the gain of the optimal filter model at a firing rate of 50 Hz (the effective bandwidths of the optimal and LGN filter models practically coincide at this firing rate, see table 2). (E), (F) Comparison of the lower bound on the rate of information transmission ($I_{LB}/2$, filled symbols) and the ϵ -entropy ($I_\epsilon/2$, empty symbols) in response to white noise (squares) and natural stimuli (circles). (E) and (F) show respectively information transmission of the optimal filtering neuron and of the LGN filter model.

Table 2. Transition from low-pass to band-pass for the optimal filter model. The columns give respectively the mean firing rate ($\lambda_0/2$) of each model neuron, the effective cut-off frequency, the frequency at which the gain $|\hat{K}(f)|$ is maximal and the ratio $\rho = |\hat{K}(f_{\text{peak}})|/|\hat{K}(0.1)|$.

$\lambda_0/2$ (Hz)	f_e (Hz)	f_{peak} (Hz)	ρ
5	3	1.3	2.2
12.5	6	2.7	3.1
25	9	4.7	4.0
50	16	8.3	5.4
75	23	11.6	6.3
100	29	14.8	7.1

The optimal filter gain encoding the Gaussian stimuli (6.5) is given by

$$|\hat{K}(\omega)|^2 = \frac{\lambda_0}{\mu} (1 + \tau^2 \omega^2) \left(\frac{\tau}{2} \frac{2\omega_0 + \pi^2 \lambda_0}{\text{arcsinh}(\tau \omega_0)} \frac{1}{(1 + \tau^2 \omega^2)^{1/2}} - 1 \right) \quad (6.6)$$

for $|\omega| < \omega_0$ and $|\hat{K}(\omega)|^2 = 0$ for $|\omega| \geq \omega_0$. The constant ω_0 entering this equation is determined numerically from the constraint (5.11). We see from equation (6.6) that μ only plays the role of an overall scaling factor so that we may set $\mu = 1$ in the following. The effective bandwidth depends on the mean firing rate of the model neuron. Table 2 gives the values of $f_e = \omega_{e2}/2\pi$ ($\omega_{e1} = 0$) for a range of firing frequencies, as determined numerically from equation (5.11). Figure 3(A) shows the modulus $|\hat{K}(\omega)|$ for different mean firing rates. The overall shape of the optimal encoding filter changes with mean firing rate making a transition from low-pass to band-pass at high firing rates. This is further quantified in table 2 which reports the frequency at which the gain is maximal and the ratio of maximal gain to gain at a fixed frequency of 0.1 Hz. The reason for this transition is clear: at low firing rates the dynamic range of the model cell is limited and low-pass filtering the signal leads to better reconstructions by eliminating signal power in a frequency band where the cell cannot encode the signal. As the mean firing rate rises, the dynamic range increases and higher stimulus frequencies can be encoded. These need, however, to be amplified with respect to low frequencies, due to the rapid decay in signal power with temporal frequency. After reaching a peak frequency, amplification becomes undesirable because it also amplifies noise. Similar effects have been studied in mean firing rate models by Atick and collaborators [3, 19, 8]. Figure 3(B) shows the mean relative errors in the reconstructions of white and natural stimuli. In both cases, the power is limited to the optimal frequency band of the model (see table 2, $f_c = 15$ Hz if $f_e < 15$ Hz). The relative mean error was computed numerically for the white stimulus (using equation (5.5)) and from equation (6.6) for the natural ensemble.

For comparison, figure 3(C) reports the relative mean error in the reconstructions of similar stimuli when the filter gain is chosen to correspond to that of an LGN cell [29] as derived in [8]. Since the power spectrum of equation (6.5) has an added low-frequency cut-off compared to the spectrum given in [8], we modified the whitening filter equation of Dong and Atick accordingly,

$$|\hat{K}(\omega)| = \frac{\eta}{\mu^{1/2}} \frac{\sqrt{1 + \tau^2 \omega^2}}{(1 + \omega^2/\omega_n^2 + 1/\tau^2 \omega_n^2)^{3/2}} \quad (6.7)$$

in order not to underestimate the encoding of low-frequency components by the model. In equation (6.7) the parameter $\omega_n = 2\pi f_n$ is taken from [8], with $f_n = 5.5$ Hz. The parameters μ and τ are those given for equation (6.5) and η is a scaling factor which

determines the mean firing rate. The effective bandwidth computed from equation (6.7) is 18 Hz ($\Delta_e = (0; 18 \text{ Hz})$) and f_c is set to this value in figure 3(C). As can be seen from the figure, at the same firing rate both filters reconstruct the natural stimulus much better than the white one. This is to be expected from the results derived in example 6.1. Furthermore, the LGN filter encodes white stimuli better than the optimal filter at low firing rates (this is due to the low cut-off frequency of the optimal filter, see table 2) but performs significantly worse than the optimal filter to encode natural stimuli. At low firing rates, its relative mean error is higher (10–20%) than that of the optimal filter and at high firing rates (50 Hz or more) the errors become comparable but its bandwidth is smaller. Indeed, the shape of the LGN gain differs from the optimal one as illustrated in figure 3(D). Nevertheless, both filters are able to encode between 10–50% of a time-varying white stimulus provided that the bandwidth is limited to 15–30 Hz. In figures 3(E) and (F), I_{LB} and I_e are plotted for these two model neurons and for the same ensembles as in figures 3(B) and (C). In both cases, the model neurons transmit more information when the stimulus is white than when it is taken from the natural ensemble, the difference being as large as 15 bit s^{-1} for firing rates of 100 Hz. In spite of this, the mean square errors in the reconstructions are always lower for the natural ensemble than for the white one (see figures 3(B) and (C)). Furthermore, the lower bound I_{LB} is slightly larger than I_e in the case of white stimuli, but the difference is pronounced for natural stimuli. In the example of the LGN filter model, I_e ranges from 0.4 to 6.5 bit s^{-1} when the firing rate varies between 5 and 100 Hz whereas I_{LB} ranges from 3.5 to 17 bit s^{-1} for the same mean firing rates.

7. Discussion

In this concluding section, we discuss in turn the neuron models used, the quantification of reconstructions and information coding in the light of the results obtained in sections 5 and 6.

7.1. The neuron models

Clearly, the model of linear and half-wave rectifying neurons with exponential threshold defined in section 4 is an idealization of biological neurons. Linearity, for example, is observed in a range of stimulus parameters, but additional compression and saturation effects usually come into play at a certain level of response [2]. While half-wave rectification has been observed to describe adequately the centre response dynamic of retinal X cells [34], in the case of simple cells an expansive nonlinearity might lead to a better description of experimental data [14]. Finally, the assumption that single spikes are generated independently of each other has to be modified if one is to take into account the experimental interspike interval statistics observed in LGN or retinal ganglion cells [18, 33]. Nevertheless, the models studied here represent a reasonable starting point to study the encoding of time-varying stimuli in single spike trains. The assumption of linear and half-wave rectification has been widely used to fit experimental data (see, e.g., [29, 7]) or in theoretical studies [8] and the additional assumption of independently generated spikes constitutes the simplest generalization of mean firing rate models which is able to take into account single action potentials. Furthermore, the resulting models have the considerable advantage of being amenable to analytical solutions, as demonstrated in the preceding sections. Finally, the connection to more general integrate-and-fire neuron models, explained in appendix B, implies that it is possible to improve our models to include all the effects mentioned above and to study them by computer simulations.

7.2. Quantification of the reconstructions

As shown in sections 2, 5 and 6, it is possible to quantify the quality of reconstructions using mean square error by first determining an ‘effective bandwidth’ and then using stimuli restricted to the effective bandwidth of the cell. The effective bandwidth defined in section 2 was shown to coincide with the usual definition in the case of linear and half-wave rectifying neurons (proposition 5.1). Hence, it plays a similar role in the reconstruction algorithm to the bandwidth defined from a first-order Wiener kernel.

The ϵ -entropy is an information theoretic quantity like I_{LB} but in contrast to it, I_ϵ is an absolute lower bound: it is not possible by any means to approximate the signal s with mean error ϵ by transmitting less information than I_ϵ . For a Gaussian white noise stimulus ensemble $I_{LB} \geq I_\epsilon$ holds (proposition 3.1) and, although no proof is available for more general ensembles, this inequality is expected to remain true (as illustrated, for example, in figures 3(E) and (F)). The quantity $I_{LB} - I_\epsilon$ represents the minimal amount of information that the neuron transmits in addition to I_ϵ to achieve a given fidelity criterion ϵ . Since it is possible to achieve the same performance by transmitting only I_ϵ , one may regard $I_{LB} - I_\epsilon$ as representing a lower bound on the ‘inefficiency’ of the neuron in reproducing the stimulus to a mean accuracy ϵ . Alternatively, we may define a ‘coefficient of coding efficiency’ $c_{\text{eff}} = I(s_{\text{est}}; s)/I_\epsilon \geq 1$, taking the value 1 if the neuron transmits exactly the minimal amount of information needed to approximate the stimulus to accuracy ϵ and $c_{\text{eff}} > 1$ if the neuron is transmitting more than the ideal minimum. It follows from proposition 3.1 that $c_{\text{eff}} \geq c_{LB} \geq 1$, where c_{LB} is defined by $c_{LB} = I_{LB}/I_\epsilon$. From properties 5.3 and 5.4 one concludes that c_{LB} will approach the theoretical lower bound of 1 in the limit of vanishing firing rates. As the mean firing rate increases, the difference between I_{LB} and I_ϵ increases as well (see figure 2(B)) and c_{LB} ranges from 1.17 ($\tau = 10$ ms) to 1.45 ($\tau = 200$ ms) for $\lambda_K/2 = 100$ Hz. Interestingly, c_{LB} depends on the choice of the stimulus ensemble, as illustrated in figures 3(E) and (F). When the stimulus is white, the value of c_{LB} for the LGN and optimal filters is not significantly different from those given above, whereas in the case of the natural stimulus ensemble c_{LB} can be as high as 1.85 (optimal filter) and 2.6 (LGN filter) for firing rates of $\lambda_K/2 = 100$ Hz. In these cases the model neurons are transmitting at least twice as much information than ideally needed to approximate the stimulus with the same accuracy. It will be of interest to compute values of c_{LB} directly from experimental data. These theoretical results are complementary to those obtained in [28]. In that study, the authors compared I_{LB} to the entropy of the spike train (computed by assuming a finite temporal resolution for the spike occurrence times) and showed that I_{LB} represents 50–60% of the available spike train entropy. This demonstrates that single cells were operating at more than half of their theoretical possibilities, but does not put the reconstruction performance on an absolute scale.

As illustrated in figures 3(E) and (F), the rate of information transmitted by model neurons, as well as the quality of reconstructions, will in general depend on the choice of the stimulus ensemble. Both I_ϵ and I_{LB} are higher for white than for natural stimuli. In contrast, the quality of reconstructions is considerably worse for the white than for the natural stimulus ensemble (see figures 3(B) and (C)).

7.3. Signal processing of linear and half-wave rectifying model neurons.

Linear and half-wave rectifying neurons are in principle able to encode arbitrarily well a time-varying Gaussian stimulus (see property 5.2; with $\mu = 0$ we obtain $\epsilon^2 \rightarrow 0$ and $I_{LB} \rightarrow \infty$ as $\lambda_K \rightarrow \infty$). However, for model parameters which are physiologically

plausible, the mean error in the estimation is not expected to be negligible (see figures 2 and 3). These models are not able to transmit more than 1.13 bit/spike of information about the stimulus, a result which is independent of the particular linear processing preceding the spiking mechanism (see properties 5.2 and 5.3). Quantities of information transmitted per spike which are compatible with this upper bound have been measured in the fly visual system [27]. On the other hand, rates of information as high as 3 bit/spike have been reported experimentally in certain preparations [28]. Assuming that our results can be extrapolated to real neurons, they would imply that some kind of nonlinear processing is underlying these high rates.

By imposing a principle of optimal computation for the encoding of signals having the statistics of natural time-varying images [8], we were able to predict a transition from low-pass filtering at low firing rates to band-pass filtering at high firing rates for the optimal encoding by a model neuron (see figure 3(A) and table 2). A transition of the same nature has been described experimentally in retinal ganglion cells [31] and similar effects were reported for LGN cells [29]. We wish to emphasize that a similar prediction could be obtained by starting from different principles of optimal computation than the one used here. For example, optimal decorrelation as applied in [8] is also expected to lead to a transition from low-pass to band-pass filtering. The interest of the derivation presented here lies in the fact that it relies on optimal encoding in single spike trains and does not use a mean firing rate model as its starting point. The optimal filter derived in example 6.2 is similar but not identical to the LGN filter obtained in [8] (see figure 3(D)). No attempt has been made here to fit the experimental data (different optimal filters could be obtained by considering different threshold statistics) and the assumption of Poisson spikes is most probably an oversimplification for LGN cells [18].

As shown in figures 2 and 3, model neurons are able to encode an appreciable portion of the time-varying changes of random stimuli in their spike trains. From these examples, one can read that approximately 20% of a white noise stimulus, having 20 Hz of bandwidth, could be recovered from a single cell. In the case of the LGN model neuron, substantially better estimates were obtained when reconstructing the natural stimulus ensemble (see figure 3). This is due to the fact that this ensemble contains substantially less power at high temporal frequencies than the white ensembles otherwise tested (see equation (6.5)). Clearly, the LGN model used is schematic and better estimates of the information effectively transmitted would require a refined analysis, but these results open the interesting possibility that single LGN neuron spike trains might be able to encode more than 50% of a time-varying natural stimulus. This hypothesis can be tested experimentally. In any case, the fact that a single model neuron is able to encode 20% of a time-varying stimulus having an approximative bandwidth of 20 Hz implies that only a small number of cells is needed to carry precise temporal information. In order to recover an estimate to 10% accuracy, averaging over $N = ((1 - 0.2)/0.1)^2 = 64$ spike trains of independent neurons will be sufficient. This estimate can be regarded as conservative since we are considering here white input signals, but we have also neglected the fact that single cells might encode a DC component in their spike trains and that there will usually be correlations between the activity of the cells. By committing spikes to encode the mean value of the signal, one expects that the capacity to register time-varying changes will be degraded. Similarly, correlations between single neurons which can be on the order of 10–20% (see [36] and references given therein) will also limit reconstruction accuracy.

Finally, we would like to emphasize that the results presented here were derived using single spike trains of ‘noisy’ neurons. In our models, spikes do not possess any particular timing precision or structure besides the modulation in firing probability induced by the

stimulus. The stimulus is reliably encoded in their mean firing rate and single spike trains always carry less information. It has recently been proposed [30], on the basis of physiological and psychophysical evidence, that information might be carried in the visual cortex of primates by fundamental signalling units consisting of pools of 100 ‘noisy’ neurons. Our results demonstrate that such an encoding scheme is theoretically possible. It remains to be explained how such a distributed system extracts and processes the information which is in principle available in single neuronal spike trains.

Acknowledgments

The author would like to thank D W Dong for sharing and explaining the results of [8] prior to publication. This work was supported by a grant of the Roche Foundation and of the National Science Foundation.

Appendix A. Equivalent formulae for the lower bound I_{LB}

Let $n_1(t)$ be the ‘noise’ in the reconstructions which is uncorrelated to the stimulus $s(t)$,

$$s_{\text{est}}(t) = (g \star (s + n_1))(t) \quad (\text{A.1})$$

where the function $g(t)$ is chosen so that

$$\langle s(t_1)n_1(t_2) \rangle = 0 \quad \forall t_1, t_2 \in \mathbb{R}. \quad (\text{A.2})$$

Equations (A.1) and (A.2) define $n_1(t)$ and $g(t)$ uniquely.

Proposition A.1. The lower bound I_{LB} can be computed using either of the following two formulae:

$$I_{LB} = \frac{-1}{4\pi \log(2)} \int_{\Delta_c} d\omega \log(1 - \hat{g}(\omega)) \quad (\text{A.3})$$

$$I_{LB} = \frac{1}{4\pi \log(2)} \int_{\Delta_c} d\omega \log \left(1 + \frac{S_{ss}(\omega)}{S_{n_1 n_1}(\omega)} \right) \quad (\text{A.4})$$

where $S_{n_1 n_1}(\omega)$ is the power spectrum of $n_1(t)$.

Remark A.1. (i) Equation (A.3) was first used by Stein *et al* [32] to estimate the rate of information transmitted by integrate-and-fire models stimulated with Gaussian white noise.

(ii) Equation (A.4) was used by Bialek and collaborators in [5, 28].

(iii) The function $\hat{g}(\omega)$ is known as the *coherence function* in system identification theory [4].

Proof. From (A.1) and (A.2) it follows that

$$\langle s(t_1)s_{\text{est}}(t_2) \rangle = (g \star R_{ss})(t_2 - t_1). \quad (\text{A.5})$$

On the other hand,

$$\langle s(t_1)s_{\text{est}}(t_2) \rangle = (h \star R_{sx})(t_2 - t_1). \quad (\text{A.6})$$

Hence, combining (A.5) and (A.6), $g \star R_{ss}(\tau) = h \star R_{sx}(\tau)$, so that after Fourier transformation, $\hat{g}(\omega)S_{ss}(\omega) = \hat{h}(\omega)S_{sx}(\omega)$, or equivalently,

$$\hat{g}(\omega) = \frac{|S_{sx}(\omega)|^2}{S_{xx}(\omega)S_{ss}(\omega)} \quad \text{for } \omega \in \Delta_c \quad (\text{A.7})$$

where we have used definition (2.3) of the filter h . It is now easy to show using equation (2.4) that

$$\frac{S_{ss}(\omega)}{S_{nn}(\omega)} = \frac{1}{1 - \hat{g}(\omega)}.$$

This completes the proof of equation (A.3).

We set $t_2 - t_1 = \tau$ and compute both sides of

$$\langle (g \star n_1)(t_1) (g \star n_1)(t_2) \rangle = \langle (s_{\text{est}}(t_1) - (g \star s)(t_1))(s_{\text{est}}(t_2) - (g \star s)(t_2)) \rangle. \quad (\text{A.8})$$

Expanding the right-hand side and computing each term separately, we obtain

$$\begin{aligned} \langle s_{\text{est}}(t_1) s_{\text{est}}(t_2) \rangle &= ((h^r \star h) \star R_{xx})(\tau) & \langle (g \star s)(t_1) s_{\text{est}}(t_2) \rangle &= ((g^r \star h) \star R_{sx})(\tau) \\ \langle s_{\text{est}}(t_1) (g \star s)(t_2) \rangle &= ((g^r \star h) \star R_{sx})(\tau) & \langle (g \star s)(t_1) (g \star s)(t_2) \rangle &= ((g \star g^r) \star R_{ss})(\tau). \end{aligned}$$

whereas for the left-hand side we have

$$\langle (g \star n_1)(t_1) (g \star n_1)(t_2) \rangle = ((g \star g^r) \star R_{n_1 n_1})(\tau)$$

where we have defined $R_{n_1 n_1}(\tau) = \langle n_1(t) n_1(t + \tau) \rangle$.

Fourier transforming (A.8), we obtain after use of equation (A.7),

$$|\hat{g}(\omega)|^2 S_{n_1 n_1}(\omega) = |\hat{h}(\omega)|^2 S_{xx}(\omega) - |\hat{g}(\omega)|^2 S_{ss}(\omega)$$

and it follows from this latter equation that

$$S_{n_1 n_1}(\omega) = \frac{1 - \hat{g}(\omega)}{\hat{g}(\omega)} S_{ss}(\omega) \quad (\text{A.9})$$

where we have again used equation (A.7). Equations (A.9) and (A.3) imply equation (A.4). \square

Appendix B. Relation between integrate-and-fire models with random threshold and Poisson models

B.1. Model description

The model of integrate-and-fire neurons which we consider was described in figure 1(C). The input signal $s(t)$ is transformed by some operator F (which represents processing of the signal prior to the spike generating mechanism) into a positive somatic current. In section 4, we assumed F to consist of linear filtering followed by half-wave rectification, with each of the positive currents $q_s^\pm(t)$ driving one model neuron. In this appendix, we consider a single neuron and denote by $q_s(t)$ the positive somatic current.

The current $q_s(t)$ is integrated to yield the somatic voltage $y(t)$. Once the membrane potential y reaches the threshold $k(t) \geq 0$, a spike is fired and the membrane voltage is reset to zero after a refractory period δ . The threshold variable $k(t)$ is constant between two spikes at times t_i and t_{i+1} , $k(t) = k_i$, $t_i < t \leq t_{i+1}$ and is updated after each spike according to a probability distribution $p(k)$. The discrete stochastic process $\{k_n\}_{n \in \mathbb{Z}}$ describing the threshold could be chosen to have more complicated properties (e.g. [12]), we will however consider only k_n 's which are independent and identically distributed random variables.

B.2. Equivalence with modulated Poisson processes

We assume that the refractory period is equal to zero ($\delta = 0$) and that the threshold is exponentially distributed, $p(k) = \mu e^{-\mu k}$. If the injected current is $q_s(t)$ we show that the output spike trains coincide with the spike trains generated by a Poisson process with mean rate $\mu q_s(t)$.

Since the output spikes are recursively determined from the threshold distribution by

$$k_n = \int_{t_{n-1}}^{t_n} dt q_s(t) \quad n = 1, 2, \dots \quad (\text{B.1})$$

the probability density that the neuron will fire action potentials at time t_1, \dots, t_n in the interval $(0; T)$ is given by

$$p(t_1, \dots, t_n) = p(k_1, \dots, k_n) \left| \det \left(\frac{\partial k_j}{\partial t_i} \right) \right| P(\rho < k_{n+1}).$$

Here $\rho = \int_0^T dt q_s(t)$, $P(\rho < k_{n+1})$ is the probability that the random variable $k_{n+1} > \rho$ and the joint probability density $p(k_1, \dots, k_n)$ for the threshold variables k_1, \dots, k_n is given by

$$p(k_1, \dots, k_n) = \mu^n \exp \left(-\mu \sum_{i=1}^n k_i \right) = \mu^n \exp \left(-\mu \int_0^{t_n} dt q_s(t) \right)$$

after use of equation (B.1). Furthermore,

$$P(\rho < k_{n+1}) = \int_{\rho}^{\infty} dk \mu e^{-\mu k} = \exp \left(-\mu \int_{t_n}^T dt q_s(t) \right).$$

Since $k_i = k_i(t_{i-1}, t_i) = \int_{t_{i-1}}^{t_i} dt q_s(t)$ we have

$$\frac{\partial k_i}{\partial t_i} = q_s(t_i) \quad \frac{\partial k_i}{\partial t_{i-1}} = -q_s(t_{i-1})$$

and

$$\det \left(\frac{\partial k_j}{\partial t_i} \right) = q_s(t_1) \cdots q_s(t_n).$$

Summing up these results,

$$p(t_1, \dots, t_n) = \exp \left(-\int_0^T dt \mu q_s(t) \right) \mu q_s(t_1) \cdots \mu q_s(t_n)$$

which is exactly the probability density of events for a Poisson process with rate $\mu q_s(t)$.

Appendix C. Proofs

Proof of proposition 3.1. (i) is obtained by first determining θ^2 from the power spectrum of the white noise and the constraint (3.3) and then solving for I_ϵ in equation (3.2). (ii) follows immediately from Jensen's inequality ([13], paragraph 6.14).

Proof of proposition 4.1. We verify only equations (4.9) and (4.10). Equation (4.8) follows in a similar manner and is easier to prove. We start from

$$\begin{aligned} & \int \text{De } x(t_a, e) x(t_b, e) P(e|s) \\ &= \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{n_1, \dots, n_i = \pm 1} \int \cdots \int_{-T}^T dt_1 \cdots dt_i x(t_a, e) x(t_b, e) Q_i(t_1, n_1; \dots, t_i, n_i | s) \end{aligned}$$

$$= e^{-\bar{v}} \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{n_1, \dots, n_i = \pm 1} \int \cdots \int_{-T}^T dt_1 \cdots dt_i \left(\sum_{k,j=1}^i n_k n_j u(t_a - t_k) u(t_b - t_j) \right) \\ \times p(t_1, n_1 | s) \cdots p(t_i, n_i | s)$$

where the left-hand side represents a convenient symbolic notation for the right-hand side. The i terms with $t_k = t_j$ in the innermost sum and the $i(i-1)$ terms with $t_k \neq t_j$ have to be treated separately,

$$\int De x(t_a, e) x(t_b, e) P(e | s) \\ = e^{-\bar{N}} \sum_{i=1}^{\infty} \frac{1}{i!} \sum_{n_1, \dots, n_i = \pm 1} \int \cdots \int_{-T}^T dt_1 \cdots dt_i \left(\sum_{j=1}^i n_j^2 u(t_a - t_j) u(t_b - t_j) \right) \\ \times p(t_1, n_1 | s) \cdots p(t_i, n_i | s) \\ + e^{-\bar{N}} \sum_{i=2}^{\infty} \frac{1}{i!} \sum_{n_1, \dots, n_i = \pm 1} \int \cdots \int_{-T}^T dt_1 \cdots dt_i \left(\sum_{\substack{j,k=1 \\ j \neq k}}^i n_k u(t_a - t_k) n_j u(t_b - t_j) \right) \\ \times p(t_1, n_1 | s) \cdots p(t_i, n_i | s)$$

and by performing first a change of integration variables and then passing the sum over n_1, \dots, n_i under the integral sign in both terms, we obtain

$$\int De x(t_a, e) x(t_b, e) P(e | s) = \int_{-T}^T dt_1 u(t_a - t_1) u(t_b - t_1) |q_s(t_1)| \\ + \iint_{-T}^{+T} dt_1 dt_2 u(t_a - t_1) u(t_b - t_2) q_s(t_1) q_s(t_2)$$

where we have used equations (4.4) and (4.5) to compute the sums over n_j and n_k , $j, k = 1, \dots, i$. It is now easy to take the $T \rightarrow \infty$ limit and to compute $\langle x(t_a) x(t_b) \rangle$ from this last equation, by averaging over the stimulus ensemble. Since the stimulus ensemble s is stationary, $\lambda_K = \langle |q_s(t_1)| \rangle_s$ is independent of t_1 and the first term of the right-hand side is seen to be equal to $\lambda_K (u^\Gamma \star u)(t_b - t_a)$. For the second term we obtain after the change of variables $t_1 \rightarrow t_1 + t_a$, $t_2 \rightarrow -t_2 + t_b$,

$$\iiint_{-\infty}^{+\infty} dt_1 dt_2 dt_3 dt_4 u(t_a - t_1) u(t_b - t_2) K(t_3) K(t_4) \langle s(t_1 - t_3) s(t_2 - t_4) \rangle_s \\ = [u^\Gamma \star (u \star (K^\Gamma \star (K \star R_{ss})))](t_b - t_a).$$

Combining these two equations we obtain the desired result.

The mean firing rate of the two neurons averaged over the stimulus ensemble can be computed by noticing that since $s(t)$ is Gaussian, the random variable $q_s(t) = \int_{-\infty}^{+\infty} dt_1 K(t_1) s(t - t_1)$, is also Gaussian for $t \in \mathbb{R}$. It has zero mean and its variance is $\sigma_{q_s}^2 = \langle |q_s(t)|^2 \rangle_s = R_{q_s q_s}(0)$, where the auto-correlation of $q_s(t)$ is given by

$$R_{q_s q_s}(\tau) = ((K^\Gamma \star K) \star R_{ss})(\tau). \quad (\text{C.1})$$

Since $q_s(t)$ is Gaussian,

$$\langle |q_s(t)| \rangle_s = 2 \int_0^{+\infty} dx \frac{1}{\sqrt{2\pi} \sigma_{q_s}} x e^{-x^2/2\sigma_{q_s}^2} = \frac{2\sigma_{q_s}}{\sqrt{2\pi}} = \frac{2}{\sqrt{2\pi}} [((K^\Gamma \star K) \star R_{ss})(0)]^{1/2}$$

using equation (C.1). Equation (4.10) is obtained by rewriting this latter expression in terms of Fourier transforms.

Proof of proposition 5.2. Equation (5.5) is obtained directly from (2.4), (2.6) and (5.3). Scaling the filter K by a positive constant η , $\hat{K}(\omega) \rightarrow \hat{K}^{(\eta)}(\omega) = \eta\hat{K}(\omega)$, $\eta > 0$, is equivalent to scaling the mean firing rate λ_K by the same constant η : $\lambda_K \rightarrow \lambda_K^{(\eta)} = \eta\lambda_K$. We therefore consider the mean square error as a function of η :

$$\epsilon^2(\eta) = \epsilon^2(K^{(\eta)}) = \frac{1}{2\pi} \int_{\Delta_c} d\omega \frac{\lambda_K S_{ss}(\omega)}{\lambda_K + \eta |\hat{K}(\omega)|^2 S_{ss}(\omega)}. \quad (\text{C.2})$$

By exchanging integral and limit (see [26], theorem I.10) we obtain from equation (C.2), $\lim_{\eta \rightarrow 0} \epsilon^2(\eta) = \sigma^2$.

In the limit $\eta \rightarrow \infty$ we have

$$\lim_{\eta \rightarrow \infty} \epsilon^2(\eta) = \int_{\Delta} d\omega S_{ss}(\omega)$$

since

$$\lim_{\eta \rightarrow \infty} \left(\frac{\lambda_K S_{ss}(\omega)}{\lambda_K + \eta |\hat{K}(\omega)|^2 S_{ss}(\omega)} \right) = \begin{cases} 0 & \text{if } \omega \in \Delta_c - \Delta \\ S_{ss}(\omega) & \text{if } \omega \in \Delta. \end{cases}$$

Finally, using the Leibnitz rule (see [6], (17.14)),

$$\frac{d}{d\eta} \epsilon^2(\eta) = \frac{1}{2\pi} \int_{\Delta_c} d\omega \frac{d}{d\eta} \left(\frac{\lambda_K S_{ss}(\omega)}{\lambda_K + \eta |\hat{K}(\omega)|^2 S_{ss}(\omega)} \right) < 0.$$

This completes the proof.

Proof of propositions 5.3 and 5.4. The proofs are completely similar to the proof of property 5.2. The inequality

$$\frac{y}{1+y} < \log(1+y) \quad y > 0$$

is needed to show that I_S is monotone decreasing and the limits (5.7) and (5.8) follow from l'Hôpital's rule and equation (4.10).

Proof of proposition 5.5. Let

$$\hat{K}_1(\omega) = \frac{1}{\lambda_K^{1/2}} \hat{K}(\omega) \quad |\hat{K}_1(\omega)|^2 = \frac{1}{\lambda_K} |\hat{K}(\omega)|^2$$

and $\tilde{\lambda}_0 = \pi^2 \lambda_0$, $X(\omega) = |\hat{K}_1(\omega)|^2$. Minimizing (5.9) under the constraint $\lambda_K = \lambda_0$ is equivalent to minimizing

$$\tilde{\epsilon}^2(X) = \int_{\Delta_c} d\omega \frac{S_{ss}(\omega)}{1 + X(\omega) S_{ss}(\omega)} \quad (\text{C.3})$$

under the constraints

$$\tilde{\lambda}(X) = \int_{\Delta_c} d\omega X(\omega) S_{ss}(\omega) = \tilde{\lambda}_0 \quad X(\omega) \geq 0. \quad (\text{C.4})$$

To find a stationary point of (C.3) under the first constraint of (C.4) we introduce a Lagrange multiplier α and the functional $F(X, \alpha)$,

$$F(X, \alpha) = \tilde{\epsilon}^2(X) + \alpha \tilde{\lambda}(X) \quad (\text{C.5})$$

$$= \int_{\Delta_c} d\omega G(X, \omega) + \alpha \int_{\Delta_c} d\omega H(X, \omega)$$

$$G(X, \omega) = \frac{S_{ss}(\omega)}{1 + X(\omega) S_{ss}(\omega)} \quad H(X, \omega) = X(\omega) S_{ss}(\omega)$$

so that the solution of the variational problem (C.5) is equivalent to solving (C.3) under the first constraint of (C.4). Solving the Euler–Lagrange equation corresponding to (C.5) for $X(\omega)$, we obtain the two possible solutions,

$$X(\omega) = \frac{\pm \alpha^{-1/2} S_{ss}(\omega)^{1/2} - 1}{S_{ss}(\omega)} \quad \omega \in \Delta_c$$

but we conclude from the second constraint of (C.4) that only the positive sign is allowed in front of $\alpha^{-1/2}$. If we let $\Delta_1 = \{\omega | \alpha^{-1/2} S_{ss}(\omega)^{1/2} > 1\} \subseteq \Delta_c$, our final solution is

$$X(\omega) = \begin{cases} \frac{\alpha^{-1/2} S_{ss}(\omega)^{1/2} - 1}{S_{ss}(\omega)} & \text{if } \omega \in \Delta_1 \\ 0 & \text{otherwise} \end{cases} \quad (\text{C.6})$$

where the Lagrange multiplier α and the domain Δ_1 have to satisfy the constraint equation $\int_{\Delta_1} d\omega S_{ss}(\omega) X(\omega) = \tilde{\lambda}_0$. Rewriting this in terms of $|\hat{K}(\omega)|^2$ we obtain proposition 5.5.

References

- [1] Abeles M, Bergman H, Margalit E and Vaadia E 1993 Spatiotemporal firing patterns in the frontal cortex of behaving monkeys *J. Neurophysiol.* **70** 1629–38
- [2] Albrecht D G and Hamilton D B 1982 Striate cortex of monkey and cat: contrast response function *J. Neurophysiol.* **48** 217–37
- [3] Atick J J and Redlich A N 1992 What does the retina know about natural scenes? *Neural Comput.* **4** 196–210
- [4] Bendat J S and Piersol A G 1986 *Random Data, Analysis and Measurement Procedures* (New York: Wiley)
- [5] Bialek W, de Ruyter van Steveninck R R and Warland D 1991 Reading a neural code *Science* **252** 1854–7
- [6] Blatter C 1979 *Analysis II* (Berlin: Springer)
- [7] DeAngelis G C, Ohzawa I and Freeman R D 1993 Spatiotemporal organization of simple-cell receptive fields in the cat striate cortex. II. Linearity of temporal and spatial summation *J. Neurophysiol.* **69** 1118–35
- [8] Dong D W and Atick J J 1994 Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus *J. Neurosci.* in press
- [9] Field D J 1987 Relations between the statistics of natural images and the response properties of cortical cells *J. Opt. Soc. Am. A* **4** 2379–94
- [10] Frishman L J, Freeman A W, Troy J B, Schwietzer-Tong D E and Enroth-Cugell C 1987 Spatiotemporal frequency responses of cat retinal ganglion cells *J. Gen. Physiol.* **89** 599–628
- [11] Gabbiani F and Koch C 1995 Coding of time-varying signals in spike trains of integrate-and-fire neurons *Neural Comput.* in press
- [12] Gestri G, Masterbroek H A K and Zaagman W H 1980 Stochastic constancy, variability and adaptation of spike generation: performance of a giant neuron in the visual system of the fly *Biol. Cybern.* **38** 31–40
- [13] Hardy G H, Littlewood J E and Pólya G 1967 *Inequalities* (Cambridge: Cambridge University Press)
- [14] Heeger D J 1992 Half-squaring in responses of cat striate cells *Vis. Neurosci.* **9** 427–43
- [15] Ihara S 1993 *Information Theory for Continuous Systems* (Singapore: World Scientific)
- [16] Kolmogorov A N 1956 On the Shannon theory of information transmission in the case of continuous signals *IRE Trans. Int. Th.* **IT-2** 102–8
- [17] Lee Y W 1960 *Statistical Theory of Communication* (New York: Wiley)
- [18] Levine M W and Troy J B 1986 The variability of the maintained discharge of cat dorsal lateral geniculate cells *J. Physiol. (Lond.)* **375** 339–59
- [19] Li Z and Atick J J 1994 Efficient stereo coding in the multiscale representation *Network: Comput. Neural Syst.* **5** 157–74
- [20] Mainen Z F and Sejnowski T J 1995 Reliability of spike generation in neocortical neurons *Science* in press
- [21] Marmarelis P Z and Marmarelis V Z 1978 *Analysis of Physiological Systems: The White Noise Approach* (New York: Plenum)
- [22] Optican L M and Richmond B J 1987 Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex *J. Neurophysiol.* **57** 162–78
- [23] Papoulis A 1991 *Probability, Random Variables, and Stochastic Processes* 3rd edn (New York: McGraw-Hill)
- [24] Poor H V 1994 *An Introduction to Signal Detection and Estimation* (New York: Springer)

- [25] Purpura K P, Victor J D and Katz E 1994 Striate cortex extracts higher-order spatial correlations from visual textures *Proc. Natl Acad. Sci. USA* **91** 8482–6
- [26] Reed M and Simon B 1980 *Methods of Modern Mathematical Physics. I: Functional Analysis* (San Diego: Academic)
- [27] Rieke F 1991 Physical principles underlying sensory processing and computation *PhD Thesis* University of California at Berkeley
- [28] Rieke F, Warland D and Bialek W 1993 Coding efficiency and information rates in sensory neurons *Europhys. Lett.* **22** 151–6
- [29] Saul A B and Humphrey A L 1990 Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus *J. Neurophysiol.* **64** 206–24
- [30] Shadlen M N and Newsome W T 1994 Noise, neural codes and cortical organization *Curr. Op. Neurobiol.* **4** 569–79
- [31] Shapley R M and Victor J D 1981 How the contrast gain control modifies the frequency responses of cat retinal ganglion cells *J. Physiol. (Lond.)* **318** 161–79
- [32] Stein R B, French A S and Holden A V 1972 The frequency response, coherence, and information capacity of two neuronal models *Biophys. J.* **12** 295–322
- [33] Troy J B and Robson J G 1992 Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance *Vis. Neurosci.* **9** 535–53
- [34] Victor J D 1987 The dynamics of the cat retinal X cell centre *J. Physiol. (Lond.)* **386** 219–46
- [35] Wiener N 1949 *Extrapolation, Interpolation and Smoothing of Stationary Time Series* (New York: Wiley)
- [36] Zohary E, Shadlen M N and Newsome W T 1994 Correlated neuronal discharge rate and its implications for psychological performance *Nature* **370** 140–3