INSTITUTE OF PHYSICS PUBLISHING NETWORK: COMPUTATION IN NEURAL SYSTEMS

Network: Comput. Neural Syst. **14** (2003) 803–824 PII: S0954-898X(03)68394-6

The effects of spontaneous activity, background noise, and the stimulus ensemble on information transfer in neurons

Maurice J Chacron[1,2,3], André Longtin[1] and Leonard Maler[2]

[1] Department of Physics, University of Ottawa, 451 Smyth Road, Ottawa, ON, K1H-8M5, Canada
[2] Department of Cellular and Molecular Medicine, University of Ottawa, 451 Smyth Road, Ottawa, ON, K1H-8M5, Canada

E-mail: mchacron@physics.uottawa.ca

Received 26 July 2002, accepted for publication 4 September 2003
Published 22 September 2003
Online at stacks.iop.org/Network/14/803

Abstract
Information theory is playing an increasingly important role in the analysis of neural data as it can precisely quantify the reliability of stimulus–response functions. Estimating the mutual information between a neural spike train and a time varying stimulus is, however, not trivial in practice and requires assumptions about the specific computations being performed by the neuron under study. Consequently, estimates of the mutual information depend on these assumptions and their validity must be ascertained in the particular physiological context in which experiments are carried out. Here we compare results obtained using different information measures that make different assumptions about the neural code (i.e. the way information is being encoded and decoded) and the stimulus ensemble (i.e. the set of stimuli that the animal can encounter in nature). Our comparisons are carried out in the context of spontaneously active neurons. However, some of our results are also applicable to neurons that are not spontaneously active. We first show conditions under which a single stimulus provides a good sample of the entire stimulus ensemble. Furthermore, we use a recently introduced information measure that is based on the spontaneous activity of the neuron rather than on the stimulus ensemble. This measure is compared to the Shannon information and it is shown that the two differ only by a constant. This constant is shown to represent the information that the neuron's spontaneous activity transmits about the fact that no stimulus is present in the animal's environment. As a consequence, the mutual information measure based on spontaneous activity is easily applied to stimuli that mimic those seen in nature, as it does not require *a priori* knowledge of the stimulus ensemble. Finally, we consider the effect of noise in the animal's environment on information transmission about sensory stimuli. Our results show that, as expected, such 'background' noise will increase the

[3] Author to whom any correspondence should be addressed.

trial-to-trial variability of the neural response to repeated presentations of a stimulus. However, the same background noise can also increase the variability of the spike train and hence can lead to increased information transfer in the presence of background noise. Our study emphasizes how different assumptions can lead to different predictions for the information transmission of a neuron. Assumptions about the computations being performed by the system under study as well as the stimulus ensemble and background noise should therefore be carefully considered when applying information theory.

1. Introduction

Information theory (Shannon 1948a, 1948b, Cover and Thomas 1991) is an increasingly popular tool to analyse input–output properties of neural systems and characterize neural coding (Borst and Theunissen 1999, Goychuk 2001). The mutual information rate, $I(R, S)$, between the neuronal response R and a stimulus S is defined as the reduction in the entropy rate of the neural response R brought about by the stimulus S. Thus, according to Shannon (1948a, 1948b), we have $I(R, S) = H(R) - H(R/S)$ where $H(R)$ is the entropy rate of the response and $H(R/S)$ is the entropy rate of the response given the stimulus S. $H(R/S)$ measures the trial-to-trial variability in the neural response that cannot be accounted for by the stimulus (Strong *et al* 1998). In an optimal situation, the trial-to-trial variability is minimized, thus minimizing $H(R/S)$, and maximizing the mutual information $I(R, S)$. On the other hand, $H(R)$ measures the uncertainty in the spike train brought about by the stimulus ensemble (i.e. the set of all stimuli that an animal can encounter in nature): it is also the maximum value that the mutual information rate $I(R, S)$ can attain. The mutual information rate is usually expressed in bits s^{-1} and expresses the capacity of a neuron to discriminate between different stimuli that are part of the stimulus ensemble per unit time. Thus, if the mutual information rate is equal to 10 bits s^{-1}, this means that the neuron is able to discriminate between 2^{10} different stimuli from 1 s of neuronal spiking activity.

There are different techniques for estimating bounds to the mutual information rate of a neuron (see Borst and Theunissen 1999, for a review) and we will now summarize them. Indirect methods of estimating information transfer such as linear reconstruction have been used to estimate $I(R, S)$ (Gabbiani 1996, Gabbiani *et al* 1996, Rieke *et al* 1996, Gabbiani and Koch 1998). These methods assume that only certain features of the stimulus are being encoded by the neuron and thus only give a lower bound estimate for I (Borst and Theunissen 1999). These methods further assume that the stimulus used has Gaussian statistics (Rieke *et al* 1995, Gabbiani 1996, Rieke *et al* 1996, Machens *et al* 2001). For these reasons, the mutual information estimate obtained is only a lower bound to the 'true' mutual information rate. However, it has been shown that natural stimuli often do not have Gaussian statistics (Simoncelli and Olshausen 2001).

Consequently, a direct method proposed by Strong *et al* (1998) that makes fewer assumptions on the nature of the neural code has been used to estimate information transfer by neurons about both artificial and naturalistic stimuli (de Ruyter van Steveninck *et al* 1997, Buracas *et al* 1998, Strong *et al* 1998, Reinagel and Reid 2000, Fairhall *et al* 2001, Lewen *et al* 2001, Nirenberg *et al* 2001). This direct method requires large amounts of data. Furthermore, it relies on the assumption that the particular stimulus 'S' used provides a good sample of the entire stimulus ensemble. However, the validity of the assumption should be tested and we do so for a subset of the stimulus ensemble that is easily characterized. Furthermore, the final

result obtained by this depends on the binwidth $\Delta\tau$ and even diverges as $\Delta\tau$ goes to zero. Thus, certain authors have proposed a different way of estimating mutual information in the context of single ion channels and rate modulated Poisson processes (Gaspard and Wang 1993, Goychuk and Hänggi 2000, Goychuk 2001). In these studies, the authors present analytical results on different information measures. Analytical treatment of complicated neural models is seldom possible. Furthermore, one must resort to numerical estimation in the case of real experimental data. We therefore choose to use the information measures proposed by Strong *et al* (1998) that unfortunately depend on the partition size but can readily be calculated from experimental data. We note that information theoretic quantities can also be computed from the interspike interval sequence (Zador 1998, Steuer *et al* 2001a, 2001b, Tiesinga 2001). However, we focus on the information carried by the entire spike train here.

This direct method (Rieke *et al* 1996, Strong *et al* 1998) of estimating information transfer has mostly been applied to neurons that are not spontaneously active. However, many neurons are known to be spontaneously active (e.g. auditory fibres (Kiang 1965, Köppl 1997), vestibular afferents (Goldberg and Fernández 1971), neocortical neurons in awake animals (Hubel 1959, Evarts 1964, Steriade 1978), as well as cerebellar deep nuclear neurons (Aizenman and Linden 1999) and Purkinje cells (Jaeger and Bauer 1994)). Furthermore, it has been proposed that these neurons must change one or more of their spontaneous discharge properties (e.g. mean, variance, autocorrelation) in order to encode a stimulus (Ratnam and Nelson 2000, Chacron *et al* 2001b, Neiman and Russell 2002).

For this reason, an information measure that uses the entropy rate of the spontaneous activity of neurons rather than that obtained from an unrepeated stimulus S' was recently introduced and applied to the electroreceptors of weakly electric fish (Chacron *et al* 2001b). This measure is identical to the information gain or Kullback entropy (Kullback 1959) between the neural response in the absence of stimulus and the one with stimulus that has been used to characterize information transfer in neurons (Neiman *et al* 1996, Goychuk and Hänggi 2000, Goychuk 2001).

We compare the information rate to the measure described above that is useful in a situation where no spontaneous activity is seen. This comparison is done in the context of a spontaneously active neuron. We also extend the measure based on spontaneous activity to situations in which a signal must be processed in the presence of ambient noise. A classic example of this is the so-called cocktail-party effect in which one must hear a single speaker in an acoustically cluttered environment (Cherry 1953).

The main conclusion of our paper is that the rate of information transmission of a neuron not only depends on the stimuli used but also on the assumptions being made about the stimulus ensemble. To illustrate this point, our study uses both stimuli with Gaussian statistics to compare with previous work as well as naturalistic stimuli with non-Gaussian statistics and is carried out with the gymnotiform weakly electric fish *Apteronotus leptorhynchus* as a model.

These fish are particularly adept at detecting prey (Nelson and MacIver 1999) and each other (Heiligenberg *et al* 1991, Zupanc and Maler 1993, Dulka *et al* 1995) using their electrosensory system. They emit a quasi-sinusoidal time varying electric field through their electric organ discharge (EOD, frequency 600–1000 Hz). P-type electroreceptors on their skin detect amplitude modulations (AMs) of this field caused by nearby objects or conspecifics (Bastian 1981, Zakon 1986). These AMs are thus the stimulus to be encoded by these neurons while the EOD is just a carrier wave. In the absence of AMs, these electroreceptors fire action potentials in response to the EOD: thus, they are thus active in the absence of a stimulus (Bastian 1981). Electroreceptor action potentials show a phase preference with respect to the EOD signal but they skip a random number of EOD cycles between each firing. Thus their activity is considered irregular despite their phase preference with respect to the EOD (CV = 0.5:

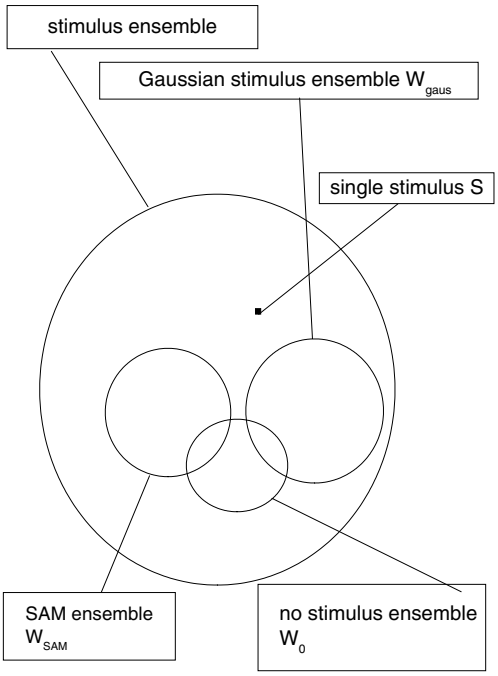

Figure 1. The stimulus ensemble W. Shown are the subset W_{gaus} of zero-mean Gaussian stimuli of contrast σ and cutoff frequency f_{c} used in this study, the subset W_{SAM} of all sinusoidal stimuli with contrast σ_{SAM} and f_{SAM}. The 'no-stimulus' ensemble W_0 is also shown; note that it intersects W_{SAM} and W_{gaus} since the cases $\sigma = \sigma_{\mathrm{SAM}} = 0$ correspond to no stimulus.

Ratnam and Nelson 2000). Consequently, there is a high trial-to-trial variability (i.e. when looking at consecutive non-overlapping epochs on spiking) in their discharge properties in the absence of stimulus (i.e. with the EOD alone). However, their activity to repeated presentations of the same AM stimulus is much more regular (Chacron *et al* 2001b).

To generate the vast amounts of data necessary for calculating the mutual information rate for such neurons, we use a simple, accurate and biophysically justified neural model that has been used to successfully account for many discharge properties of P-type electroreceptors (Chacron *et al* 2000, 2001a, 2001b). A full description and biophysical justification of the model can be found in Chacron *et al* (2001b). We also used the same parameter values as in that study.

2. Materials and methods

We first describe in detail the concept of the stimulus ensemble as well as several subsets of interest. We then discuss the methodology used to calculate the noise entropy rate. Finally, we present the different information measures we will use as well as the way they were estimated.

2.1. The stimulus ensemble

We use both artificial and naturalistic stimuli in this study. Both are part of the total stimulus ensemble denoted by W which consists of all stimuli S that the animal can encounter and is illustrated in figure 1. Each member S of the stimulus ensemble W has an associated

probability of occurrence $P(S)$. In some cases, these probabilities have been experimentally measured such as for light contrast in the fly visual system (Laughlin 1981). However, they are presently unknown for our system. The stimulus ensemble W is in general multi-dimensional and is hard to fully describe. For these reasons, researchers often concentrate on subsets that are easier to characterize.

One such subset is the ensemble of filtered Gaussian white noise stimuli that has been widely used in characterizing the information transfer properties of neurons (Gabbiani *et al* 1996, Roddey and Jacobs 1996, Wessel *et al* 1996, Clague *et al* 1997, Warland *et al* 1997). We denote this ensemble by W_{gaus}. Each member of W_{gaus} is characterized by two parameters: its contrast σ and its cutoff frequency f_{c}. The contrast is defined as the standard deviation-to-mean ratio. All Gaussian stimuli used have a mean equal to the baseline EOD amplitude value and this value is set to 1 mV. Furthermore, all Gaussian stimuli have a power spectrum that is constant and positive for frequencies between zero and the f_{c} and zero otherwise (Wessel *et al* 1996). Since every member of W_{gaus} is uniquely characterized by the parameters σ and f_{c}, the set W_{gaus} is two-dimensional. It is thus easy to fully characterize but, as before mentioned, natural stimuli seldom display Gaussian statistics.

We now describe a subset of natural stimuli that is easily characterized for our system. When two weakly electric fish with different EOD frequencies are in proximity of one another, there is interference producing a beating phenomenon, between the quasi-sinusoidal electric fields generated by each fish. This interference will cause a sinusoidal amplitude modulation (SAM) in each fish's own electric field (Bastian 1981, Heiligenberg *et al* 1991). The frequency of this modulation depends on the frequency difference between the two fish and can range from a few hertz to 400 Hz. We thus used SAMs within the physiological range as natural stimuli: other natural electrocommunication signals such as chirps (Zupanc and Maler 1993) are not considered here. We denote by W_{SAM} the set of all such stimuli. Each member of this set can be described by two parameters: the SAM contrast σ_{SAM} and frequency f_{SAM} of the sinusoid. However, to fully characterize the set W_{SAM}, we would need to know the probability of natural occurrence of a SAM with contrast σ_{SAM} and frequency f_{SAM} in nature for the fish (this would be an experiment similar to the one performed by Laughlin (1981) for the fly). Although such data can be acquired, this has to our knowledge not yet been accomplished. Note that W_{SAM} is but a small subset of W as the latter contains other natural stimuli such as the chirps mentioned earlier as well as stimuli caused by prey (Nelson and MacIver 1999).

Finally, we discuss the notion of a 'no-stimulus' ensemble. We denote this set by W_0 because this set contains only one member: a constant stimulus whose value is equal to the baseline EOD amplitude (i.e. no stimulus). It is easily seen that this set contains all SAM and Gaussian stimuli with zero contrast. As such, W_0 intersects both W_{gaus} and W_{SAM} (figure 1).

2.2. *Estimating the mutual information rate*

The calculation of the mutual information rate $I(R, S)$ requires the calculation of both the response entropy rate averaged over the stimulus ensemble $H(R)$ and the response entropy rate given the stimulus $H(R/S)$. The estimation of $H(R/S)$ is standard (Strong *et al* 1998) and poses no conceptual difficulty. We will describe it first.

2.3. *Estimating the entropy rate given the stimulus*

Our method is the same as the one used by Strong *et al* (1998). We assume that the same stimulus S was repeated many times under identical conditions and that each trial resulted in a particular spike train representing the neural response R to a particular trial. As before

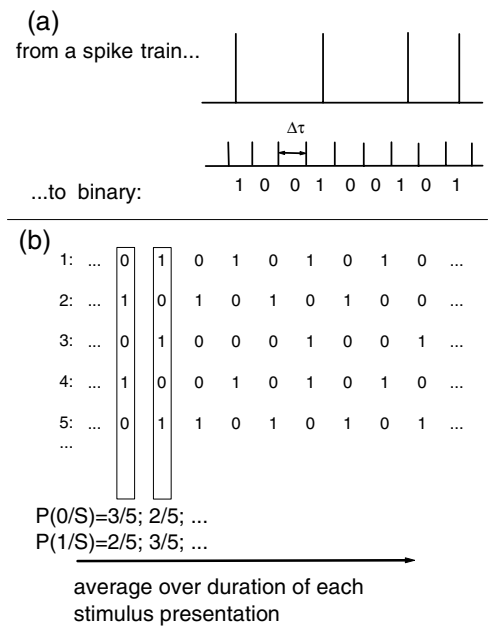

Figure 2. (a) Outline of the procedure for turning a spike train into a binary string consisting of 0 and 1. The spikes are represented by vertical bars while time is represented horizontally. The time axis is then divided into bins of length $\Delta\tau$ and each bin is assigned the value 1 if a spike occurred and zero otherwise. The binary string obtained will of course depend on the bin width $\Delta\tau$. (b) Outline of the procedure used to calculate the probabilities of occurrence of words w when the length of these words is one (i.e. 0 and 1). Multiple stimulus presentations will give rise to multiple spike trains that can be converted into binary data using the procedure outlined above. The probabilities of occurrence of words of length one are calculated at each bin from all trials. The final probabilities of occurrence $P(w/S)$ are calculated by averaging the results obtained for each bin. This procedure is then repeated for words of length two, etc. Note that the windows used for calculating the probabilities are non-overlapping.

mentioned, the response entropy rate given the stimulus $H(R/S)$ measures the response trial-to-trial variability to a stimulus.

To estimate $H(R/S)$, we partitioned the spike train into bins of length $\Delta\tau$ with $\Delta\tau$ smaller than the minimum interspike interval so that each bin contains at most one spike. A bin is then assigned the value 0 if no spikes occurred during that bin and 1 otherwise. This procedure converts a spike train into a binary string and is illustrated in figure 2(a). The alphabet thus consists of two symbols (0: no spike, and 1: spike) and words consist of binary strings of various lengths (e.g. '010', '00101').

The choice of $\Delta\tau$ is critical and depends on various considerations such as the spike timing jitter of the particular neuron (i.e. the degree to which the precise timing of action potentials is important for information transfer). This spike timing jitter can usually be measured by adding 'jitter' (e.g. in the form of Gaussian random variable) to the spike times and looking at the resulting loss in information (Kreiman *et al* 2000). On the other hand, one can also measure it by looking at the precision of spike timing (Mainen and Sejnowski 1995). Another important factor is the temporal discrimination of postsynaptic neurons (Bender *et al* 2001): this usually depends on the properties of the synapse(s) and possible inter-neuron(s) connecting the neurons in question as well as on the temporal integration properties of the post-synaptic neurons themselves. In our case, the EOD cycle is a natural time scale since P-receptor afferents

can fire at most one spike per EOD cycle (Bastian 1981). It is also known that electroreceptor input onto basilar pyramidal cells results in excitatory postsynaptic potentials whose duration is barely greater than an EOD cycle (Berman and Maler 1998). We thus take $\Delta\tau$ to be around one EOD cycle. Since we assume an EOD frequency equal to 1000 Hz, one EOD cycle corresponds to 1 ms.

We mark a particular time t with respect to the stimulus and compute the probabilities of occurrence across trials of each word ω of length L that begin at t. These probabilities are denoted by $P(w/S, t)$. The entropy of words of length L is then calculated as (Strong *et al* 1998)

$$H(L/S, t) = \sum_{w \in \Gamma(L)} -P(w|S, t) \log_2 P(w|S, t) \tag{1}$$

where $\Gamma(L)$ is the set of words of length L (e.g. if $L = 2$, then $\Gamma(2) = \{01, 10, 00, 11\}$). The entropy $H(L/S, t)$ thus denotes the local entropy of words of length L in the time window $(t, t + L)$. These entropies $H(L/S, t)$ are then averaged across all such windows during the entire time course of the stimulus. We thus have

$$H(L/S) = \langle H(L/S, t) \rangle_t \tag{2}$$

where the angle brackets denote the average over the stimulus time course. This procedure is illustrated graphically in figure 2(b).

The entropy rate of the spike train given the stimulus is then obtained by Strong *et al* (1998):

$$H(R/S) = \lim_{L \to \infty} \frac{H(L/S)}{L\,\Delta\tau}. \tag{3}$$

In practice, the entropies $H(L/S)$ calculated from equation (2) underestimate the real value for large word length L because of undersampling due to the finiteness of the spike train data. We hence used an extrapolation procedure from Strong *et al* (1998) to get an estimate of the entropy rate (Chacron *et al* 2001b):

$$\frac{H(L/S)}{L\,\Delta\tau} = H(R/S) + \frac{C_1}{L} + \frac{C_2}{L^2} + \cdots \tag{4}$$

where C_1, C_2 are constants. To estimate the entropy rate $H(R/S)$, we plot the quantities $H(L)/(L\Delta\tau)$ as a function of $1/L$ and perform either a linear or quadratic least-squares fit. The entropy rate $H(R/S)$ is then estimated from the value of the fitted line or parabola at $1/L = 0$. An example is given in figure 3. Sometimes, a quadratic fit was better than a linear one (figure 3(a)). Although we have not verified this directly, we think that the curvature is due to the ISI correlations displayed by the model (Chacron *et al* 2001b) that are present experimentally (Chacron *et al* 2000). When noise is added to the model, the ISI correlations disappear (not shown) and the results obtained from a linear fit are not very different than those obtained from a quadratic fit (figure 3(b)). Thus, linear and quadratic fits can give very different results for the entropy rates.

We estimated the entropy rate $H(R/S)$ from 1000 trials each lasting 10 s. Note that such numbers are necessary due to the high firing rates of the electroreceptors we study. These usually range from 100 to 600 Hz (Ratnam and Nelson 2000). This as well as other considerations (see above) forces us to take a small bin size $\Delta\tau$. However, for other systems, a larger $\Delta\tau$ can sometimes be taken if the firing rate is lower, and as few as 100 repetitions can sometimes lead to a good estimate of the response entropy rate given the stimulus (Strong *et al* 1998, Reinagel and Reid 2000).

Having discussed our estimation procedure for the response entropy rate $H(R/S)$, we can now discuss estimation procedures for the entropy rate of the response averaged over the stimulus ensemble. However, we first start by stating the theoretical definition.

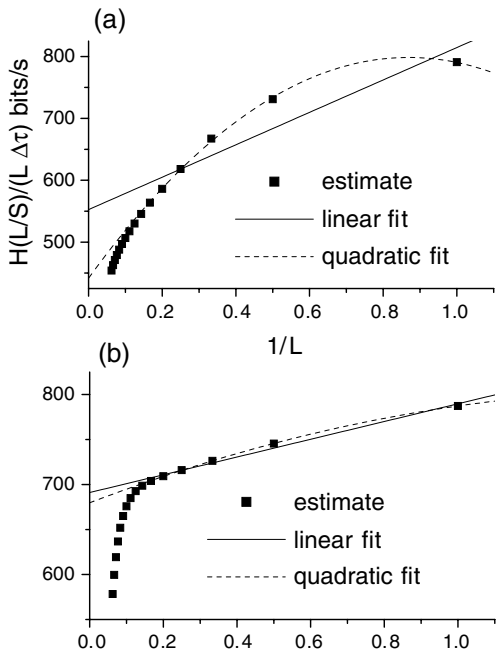

Figure 3. The entropies $H(L/S)/(L\,\Delta\tau)$ as a function of $1/L$. (a) Estimates obtained for the baseline entropy rate H_{spon} (i.e. $\sigma = 0$ mV). It is seen that a quadratic fit is much better than a linear one. (b) Estimates obtained for the baseline entropy rate in the presence of background noise for $\sigma = 0$ mV. The estimates obtained from a linear and quadratic fit are much closer in this case. We used $\Delta\tau = 1.25$ ms.

2.4. Definition of Shannon's mutual information rate

As before mentioned, the mutual information rate $I(R, S)$ is the difference between the response entropy rate averaged over the stimulus ensemble $H(R)$ and the response entropy rate given the stimulus $H(R/S)$. According to probability theory, we have that

$$P(w, t) = \sum_{S \in W(S)} P(w|S, t)P(S) \tag{5}$$

where $P(S)$ is the probability of occurrence of stimulus S within the stimulus ensemble W and $P(w/S, t)$ is, as before, the probability of w occurring at time t with respect to the stimulus. The entropy rate $H(R)$ could then be estimated by using equations (1)–(4) with $P(w, t)$ instead of $P(w/S, t)$ in equation (1) and $H(R)$ instead of $H(R/S)$ in equations (3) and (4). However, as mentioned above, the stimulus ensemble W cannot be characterized fully in general and the probabilities of occurrence of various stimuli $P(S)$ are not known. It thus becomes necessary to make simplifying assumptions at this point and we now discuss some of them.

2.5. Response entropy rate calculated over the Gaussian stimulus ensemble

It is possible to restrict oneself to an easily characterized subset of the stimulus ensemble W where the calculation described in the previous paragraph can be undertaken. Such a subset is the Gaussian stimulus ensemble W_{gaus}. We shall denote the estimate obtained by $H_{\text{gaus}}(R)$ to emphasize its dependence on the subset W_{gaus}.

As such, we estimated the entropy rate $H_{\text{gaus}}(R)$ in the following way. Gaussian stimuli of various contrasts σ and cutoff frequencies f_c each lasting 10 s were each given 1000 times. The probabilities $P(w/S, t)$ were estimated for each Gaussian stimulus S as for the calculation of the response entropy rate given the stimulus S. These probabilities were then averaged using equation (5) with W_{gaus} instead of W to obtain the probabilities $P(w, t)$. It was assumed that all stimuli were equiprobable. Equations (1)–(4) with $P(w, t)$ instead of $P(w/S, t)$ in equation (1) and $H_{\text{gaus}}(R)$ instead of $H(R/S)$ in equations (3) and (4) were then used to compute $H_{\text{gaus}}(R)$. The mutual information rate can thus be estimated as $I_{\text{gaus}}(R, S) = I_{\text{gaus}} = H_{\text{gaus}}(R) - H(R/S)$.

2.6. *Mutual information rate calculated from an unrepeated stimulus S*

Even averaging over the Gaussian stimulus ensemble can require much more data than are available. Thus, Strong *et al* (1998) have proposed that the response entropy rate $H(R)$ could be estimated from the response entropy rate of the spike train in the presence of an unrepeated stimulus S. This is based on the assumption that a single member S of the stimulus ensemble W would provide a good sample of the full stimulus ensemble W. The spike train entropy rate estimated in this way will of course depends on the particular stimulus S used and we shall denote it by $H_{\text{strong}}(R)$. The corresponding mutual information rate is thus $I_{\text{strong}}(R, S) = I_{\text{strong}} = H_{\text{strong}}(R) - H(R/S)$.

We presented a long (10 000 s duration) Gaussian stimulus and obtained a spike train that was converted into a binary sequence as described above. We estimated the probabilities of occurrence $P_{\text{stim}}(w)$ of words w within the sequence in the following way. The binary sequence was divided into non-overlapping windows of length L. The probabilities of occurrence of words of length L were then estimated across these time windows. This allowed us to compute the entropies $H_{\text{strong}}(L)$ as

$$H_{\text{stim}}(L) = -\sum_{w \in \Gamma(L)} P_{\text{stim}}(w) \log_2 P_{\text{stim}}(w). \tag{6}$$

The entropy rate $H_{\text{stim}}(R)$ was then estimated from

$$\frac{H_{\text{stim}}(L)}{L\,\Delta\tau} = H_{\text{stim}}(R) + \frac{C_1}{L} + \frac{C_2}{L^2} + \cdots. \tag{7}$$

The mutual information rate can then be estimated as $I_{\text{strong}}(R, S) = I_{\text{strong}} = H_{\text{strong}}(R) - H(R/S)$. However, this estimate assumes that the stimulus used to compute $H_{\text{strong}}(R)$ provides a good sample of the stimulus ensemble W.

2.7. *Mutual information calculated from the spontaneous activity of the neuron*

We now present an information estimate that does not depend on the stimulus ensemble *per se*. It is known that many neurons are spontaneously active and that, in most cases, this spontaneous activity *in vivo* is irregular. As such, we have previously used an information measure based on the entropy rate of the spontaneous activity of such neurons $H_{\text{spon}}(R)$ (Chacron *et al* 2001b). To estimate it, one takes a long record (10 000 s in our case) of spontaneous activity that is then converted into a binary sequence. We then used the same estimation procedure as for $H_{\text{strong}}(R)$ except that no stimulus was given.

The difference between the spontaneous response entropy rate $H_{\text{spon}}(R)$ and the response entropy rate given the stimulus $H(R/S)$ is a measure of information (Kullback 1959). We shall denote this quantity by $I_{\text{spon}}(R) = I_{\text{spon}} = H_{\text{spon}}(R) - H(R/S)$. However, this quantity is not a measure of information in the 'classical' sense since it does not take into account the stimulus

ensemble W. Instead, this quantity measures the reduction of the spontaneous response entropy brought about by a stimulus S. This is based on the assumption that spontaneously active neurons transmit information about stimuli through a change from their spontaneous activity (Ratnam and Nelson 2000, Chacron *et al* 2001b). This measure assumes that the 'no-stimulus' ensemble provides a good sample of the full stimulus ensemble W. We now discuss the relation between I_{spon} and the Shannon mutual information rate I.

We have

$$\begin{aligned} I(R, S) &= H(R) - H(R/S) \\ &= H(R) - H_{\mathrm{spon}}(R) + H_{\mathrm{spon}}(R) - H(R/S) \\ &= H(R) - H_{\mathrm{spon}}(R) + I_{\mathrm{spon}} \\ &= C + I_{\mathrm{spon}}. \end{aligned} \tag{8}$$

Thus, the mutual information rate calculated from the spontaneous activity of the neuron I_{spon} differs by only a constant C from Shannon's mutual information rate $I(R, S)$. This constant C depends on the stimulus ensemble W. Let us assume that no stimulus is given to the neuron, we then have $I(R, S) = C$. Thus C represents the information rate that a spontaneously active neuron transmits about the fact that no stimulus is present in the animal's environment and depends only on the stimulus ensemble and on the neuron's intrinsic properties. It is thus expected that I_{spon} and $I(R, S)$ will behave in the same manner when the stimulus S is varied.

Finally, note that all these measures are based on the assumption of a noise-free environment in which stimulus encoding and decoding is performed.

2.8. Noise in an animal's environment

We consider the effects of background noise on information transfer. Common sources of background noise in our system include environmental factors and conspecifics. For example, lightning is common in these fish's natural habitat and generates electrical pulses and oscillations that propagate over hundreds of kilometres through tropical waters (Hopkins 1973). The lightning pulses interfere with weakly electric fish's abilities to detect AMs (i.e. stimuli). Moreover, the fish must detect prey (invertebrates) that may be hiding in root masses (Crampton 1998). The electric image caused by these root masses can be considered noise when the fish is trying to detect its prey. Finally, two fish might be able to detect each other even when other fish are close by (Parthridge and Heiligenberg 1980): this is the weakly electric fish equivalent of the cocktail-party effect discussed above. From the above examples, it is clear that background noise will be very system dependent. In our study, we assume background noise to be Gaussian with contrast $\sigma = 0.01$ mV and cutoff frequency $f_{\mathrm{c}} = 50$ Hz. These parameters are fixed throughout. Furthermore, we will assume that the background noise is additive (i.e. the noise is added to the stimulus).

2.9. Mutual information rate in the presence of noise in the environment

Often, sensory neurons must thus encode and decode stimuli that are perturbed by variable amounts of noise such as in the cocktail party effect discussed above. We now extend the measure I_{spon} to include the effects of background noise. To do so, we must consider the effects of background noise on both the spontaneous response entropy rate $H_{\mathrm{spon}}(R)$ and on the response entropy rate given the stimulus $H(R/S)$.

We computed the response entropy rate $H_{\mathrm{noise}}(R)$ in the same manner as $H_{\mathrm{strong}}(R)$ except that background noise was given to the neuron instead of a stimulus. We computed the response entropy rate given the stimulus in the presence of background noise in the

Table 1. Summary of the different information measures.

Information measure	Description	Stimulus ensemble used	Assumptions
I_{gaus}	Mutual information rate calculated over the Gaussian stimulus ensemble W'	The Gaussian stimulus ensemble W_{gaus}	The Gaussian stimulus ensemble provides a good sample of the total stimulus ensemble W
I_{strong}	Mutual information rate calculated from a single member S of the stimulus ensemble	W	The stimulus ensemble is well sampled by the single member S
I_{spon}	Mutual information rate calculated from the spontaneous activity of a neuron	The 'no-stimulus' ensemble W_0	The neuron must have sufficiently irregular spontaneously activity
I_{noise}	Mutual information rate calculated in the presence of background noise	The 'no-stimulus' ensemble in the presence of background noise	The background noise must elicit sufficiently irregular activity from the neuron
I	In theory: mutual information rate calculated from the stimulus ensemble W	The full stimulus ensemble W	In theory: none except that the stimulus ensemble W is properly characterized

same manner as the response entropy rate given the stimulus $H(R/S)$. The only difference is that background noise is added to the stimulus before each trial. We will denote the estimate obtained by $H_{\text{noise}}(R/S)$. The mutual information rate can then be estimated as $I_{\text{noise}}(R, S) = I_{\text{noise}} = H_{\text{noise}}(R) - H_{\text{noise}}(R/S)$.

Table 1 summarizes the different information rate estimates we have presented, their assumptions, and a brief description on their estimation procedure. We will now compare results obtained using these different estimates.

3. Results

We first present results of calculations of the mutual information rate measures I_{strong}, I_{gaus}, I_{spon} and I_{noise} using Gaussian stimuli (i.e. the stimuli are part of W_{gaus}) of various physiologically relevant contrasts σ and cutoff frequencies f_c. The estimates I_{strong}, I_{spon} and I_{noise} are then compared for naturalistic SAMs of various frequencies.

3.1. The effects of varying the contrast σ

We wish to study the effects of varying the stimulus contrast σ. We thus first keep the cutoff frequency constant at 100 Hz and compare the different mutual information rate estimates obtained by varying the stimulus contrast σ. The results are shown in figure 4(a). We see that the estimates I_{gaus} (the mutual information rate estimated over the Gaussian stimulus ensemble W_{gaus}) and I_{strong} (the mutual information rate estimated from a single stimulus of W_{gaus}) are in good agreement for high contrast σ. However, I_{strong} is less than I_{gaus} as σ tends to zero. To understand this, let us assume that the stimulus has $\sigma = 0$. We then have that the entropy rate $H(R/S)$ is equal to the response entropy rate $H_{\text{strong}}(R)$, which in turn is equal to the spontaneous response entropy rate $H_{\text{spon}}(R)$. As a consequence, the estimate $I_{\text{strong}} = H_{\text{strong}}(R) - H(R/S)$ is equal to zero. However, the response entropy rate averaged over the Gaussian stimulus ensemble $H_{\text{gaus}}(R)$ is not equal to the response entropy rate $H(R/S)$ when the stimulus S has zero contrast. Hence, we have $I_{\text{gaus}} = H_{\text{gaus}}(R) - H_{\text{spon}}(R) > 0$ when $\sigma = 0$.

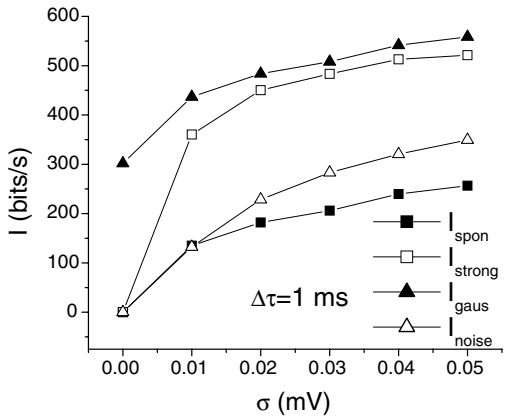

Figure 4. Effects of varying the contrast σ. Note the overall decrease in mutual information as $\Delta\tau$ increases. I_{gaus} and I_{strong} are much higher than I_{spon} and we have $I_{\text{gaus}} \approx I_{\text{strong}}$ for high contrasts. We used $\Delta\tau = 1$ ms and $f_{\text{c}} = 100$ Hz. Furthermore, we performed quadratic fits.

Our results thus show that approximating the Gaussian stimulus ensemble averaged response entropy rate $H_{\text{gaus}}(R)$ by the response entropy rate obtained from a single member of the stimulus ensemble $H_{\text{strong}}(R)$ is justified for mid to high contrast σ at this particular cutoff frequency f_{c}. This is due to the particular nature of the stimulus ensemble W_{gaus}. Note that a stimulus S with contrast σ and cutoff frequency f_{c} contains all Gaussian stimuli with lower contrasts and cutoff frequency. Thus, a member with a high contrast and cutoff frequency provides a good sample of the stimulus ensemble W_{gaus} and thus provides a good estimate of the spike train entropy rate calculated over this subset W_{gaus} of the stimulus ensemble W.

It is seen that the mutual information rate estimated over the Gaussian stimulus ensemble I_{gaus} and the mutual information rate estimated over the 'no-stimulus' ensemble I_{spon} differ by a constant. This is expected from the previous discussion since $I_{\text{gaus}} = I_{\text{spon}} + C'$ with $C' = H_{\text{gaus}}(R) - H_{\text{spon}}(R)$. The constant C' measures how well the neuron can discriminate the absence of a stimulus from other members of the stimulus ensemble W_{gaus}. This extra information is absent from the estimate I_{gaus} when contrast $\sigma = 0$ because the entropy rate $H_{\text{gaus}}(R)$ is estimated from the same stimulus ensemble member S (the 'no-stimulus' case) that is used to estimate the noise entropy rate $H(R/S)$.

Note that the mutual information rate estimated from the spontaneous activity I_{spon} is equal to 0 when the stimulus contrast $\sigma = 0$ (i.e. no stimulus is present), hence this measure implies that no information is transmitted by the neuron when no stimulus is present whereas the measure I_{gaus} implies that information is being transmitted when no stimulus is present (i.e. about the absence of stimulus). We will return to this point in the discussion.

Finally, we note that the mutual information rate calculated with background noise I_{noise} is slightly greater than I_{spon} for high contrasts but is still lower than I_{gaus}. To understand this, note that the spike train entropy rate in the presence of background noise $H_{\text{noise}}(R)$ is always greater than the spike train entropy rate in the absence of background noise $H_{\text{spon}}(R)$ (i.e. with spontaneous activity only) because the extra randomness due to background noise increases the entropy rate. For the same reason, we have that the response entropy rate given the stimulus in the presence of background noise $H_{\text{noise}}(R/S)$ is always greater than the noise entropy rate in the absence of background noise $H(R/S)$. As such, the signal-to-noise ratio (SNR) (i.e. the ratio of the stimulus contrast to the noise contrast) increases with stimulus contrast and consequently the noise becomes less and less important. The noise entropy rate

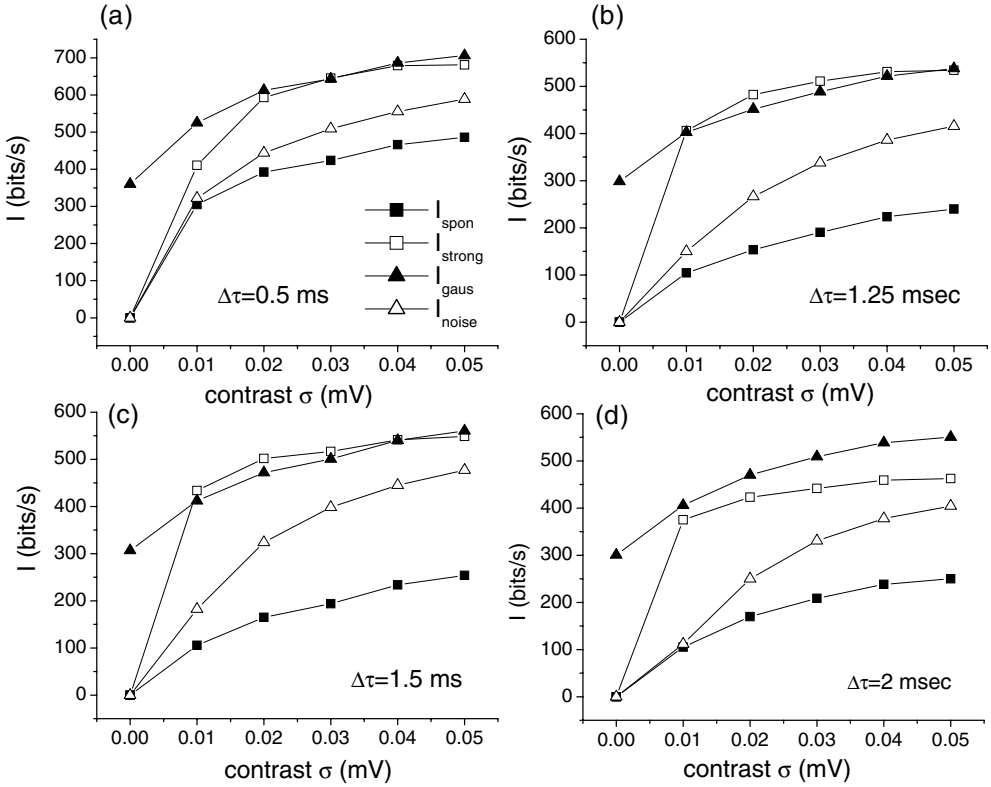

Figure 5. Effects of varying the binwidth $\Delta\tau$. Shown are the information measures I_{spon}, I_{noise}, I_{strong}, I_{gaus} as a function of stimulus contrast σ for $f_c = 100$ Hz for $\Delta\tau = 0.5$ ms (a), $\Delta\tau = 1.25$ ms (b), $\Delta\tau = 1.5$ ms (c), $\Delta\tau = 2$ ms (d). Quadratic fits were performed.

in the presence of background noise $H_{\text{noise}}(R/S)$ tends towards the noise entropy rate obtained when no background noise is present $H(R/S)$, hence making the mutual information rate in the presence of background noise I_{noise} higher than I_{spon}.

Note also that I_{noise} is less than or equal to the mutual information calculated from a single member of the stimulus ensemble I_{strong}: this is because the noise entropy rate in the presence of background noise $H_{\text{noise}}(R/S)$ is higher than the noise entropy rate obtained in the absence of background noise $H(R/S)$ thereby making the mutual information rate obtained in the presence of background noise I_{noise} lower. Moreover, the spike train entropy rate obtained from a single member of stimulus ensemble $H_{\text{strong}}(R)$ is higher than the spike train entropy rate $H_{\text{noise}}(R)$ if the stimulus S has higher contrast σ or cutoff frequency f_c than the background noise (note that the two are equal if the stimulus and background noise have the same contrast σ and cutoff frequency f_c).

3.2. Effects of varying the binwidth $\Delta\tau$

The information measures calculated in this paper depend of course on the binwidth $\Delta\tau$. In this section, we show the effects of varying the binwidth $\Delta\tau$ on the results presented in the previous section. Figures 5(a)–(d) shows the measures I_{strong}, I_{gaus}, I_{spon} and I_{noise} as a function of stimulus contrast σ for $\Delta\tau = 0.5$, 1.25, 1.5 and 2 ms, respectively. As expected, there is

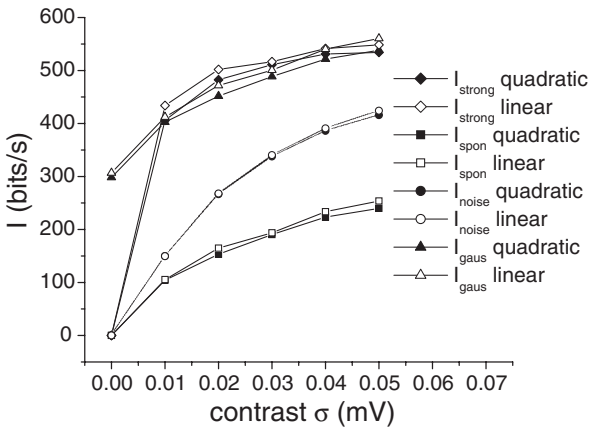

Figure 6. Effects of linear versus quadratic fits. No difference is seen in the estimates qualitatively and the quantitative variations are less than 7%. We used $\Delta\tau = 1.25$ ms.

an increase in mutual information rate as $\Delta\tau$ decreases. However, there are no qualitative differences with the case $\Delta\tau = 1$ ms presented in figure 4. Our results thus are robust to a fourfold variation in $\Delta\tau$.

3.3. Effects of linear and quadratic fits

We now show the effects of performing linear versus quadratic fits of the quantities $H(L/S)/(L\Delta\tau)$ versus $1/L$ as discussed previously (see figure 3). Figure 6 shows the information measures I_{strong}, I_{gaus}, I_{spon} and I_{noise} as a function of stimulus contrast σ for $\Delta\tau = 1.25$ ms obtained from quadratic (full symbols) and linear (open symbols) fits of the entropy rates. Essentially no qualitative difference is observed. Furthermore, there is a less than 7% difference in each measure computed from the quadratic fits versus linear fits. Similar results were observed for other $\Delta\tau$ values (not shown). Our results are thus robust to either quadratic or linear fits. As our results are also robust to changes in $\Delta\tau$, we only present results obtained with $\Delta\tau = 1$ ms and using quadratic fits from now on.

3.4. Gaussian stimuli: effects of varying the cutoff frequency f_c

We now keep the contrast σ constant at 0.03 mV and vary the cutoff frequency f_c of the stimulus. Results are shown in figure 7. We note that all four measures increase with f_c. This increase is due to the high-pass filtering characteristics of the electroreceptors we study (Xu *et al* 1996) and which are incorporated in the model we use (Chacron *et al* 2001b, Chacron 2003). The mutual information rate calculated from the Gaussian stimulus ensemble I_{gaus} and the mutual information rate calculated from a single member of the stimulus ensemble I_{strong} agree for high cutoff frequencies f_c. This is again due to a particular property of the Gaussian stimulus subset W_{gaus}. Because the power spectrum of a stimulus with cutoff frequency f_c contains all frequencies up to f_c, the stimulus essentially contains all stimuli with the same contrast σ with lower cutoff frequencies f_c. For this reason, a stimulus with a high cutoff frequency and a reasonably high contrast σ provides a good sample of the entire stimulus subset W_{gaus}. However, our results show that a single Gaussian stimulus with low frequency content will not provide a good sample of the stimulus ensemble subset W_{gaus}; consequently, the mutual information rate transmitted by the neuron will be underestimated if one uses I_{strong}.

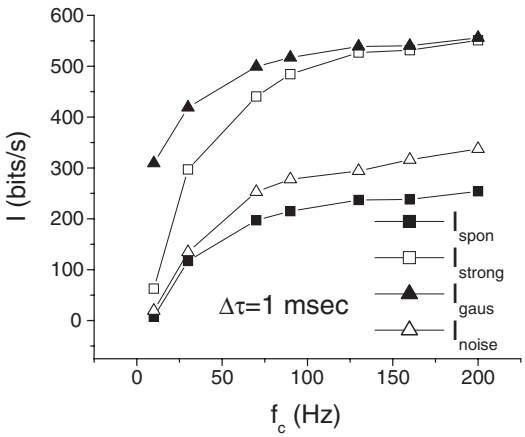

Figure 7. Information rate measures as a function of cutoff frequency for $\sigma = 0.03$ mV. Note again that $I_{\text{gaus}} \approx I_{\text{strong}}$ for high cutoff frequencies.

Note that the mutual information rate calculated over the Gaussian stimulus ensemble I_{gaus} again tends towards the constant $H_{\text{gaus}}(R) - H_{\text{spon}}(R)$ as f_c goes to zero.

The mutual information rate calculated from the spontaneous activity I_{spon} is again lower than the mutual information rate calculated over the Gaussian stimulus ensemble I_{gaus} for the same reason as above: the measure I_{spon} only gives us the reduction in the spike train entropy rate based on the 'no-stimulus' ensemble (i.e. spontaneous activity) rather than the reduction in entropy rate from the Gaussian stimulus ensemble W_{gaus}. Further, the mutual information rate obtained in the presence of background noise I_{noise} is lower than the mutual information rate I_{gaus}. This is again for the same reasons as above. However, the mutual information rate in the presence of background noise I_{N} is again higher than the mutual information rate in the absence of background noise I_{spon} since the spike train entropy rate on the presence of background noise $H_{\text{noise}}(R)$ is always greater than or equal to the spike train entropy rate in the absence of background noise $H_{\text{spon}}(R)$ with equality only in the limit where the background noise goes to zero. Due to the aforementioned high-pass filtering characteristics of the neurons we study, increasing the cutoff frequency of the stimulus f_c increases the SNR, and thus the noise entropy rate in the presence of background noise $H_{\text{noise}}(R/S)$ tends towards the noise entropy rate in the absence of background noise $H(R/S)$ because the noise becomes less and less important.

3.5. Naturalistic stimuli

As before mentioned, the mutual information rates calculated from the spontaneous activity of the spike train in the absence (I_{spon}) and presence (I_{noise}) of background noise easily extend to naturalistic stimuli as they do not require *a priori* knowledge of the neuron's full stimulus ensemble W. Figure 8(a) shows the measures I_{spon} and I_{noise} obtained by varying the SAM contrast σ_{SAM} and for constant SAM frequency $f_{\text{SAM}} = 100$ Hz. We see that in the absence of background noise, the mutual information rate I_{spon} increases with SAM contrast. This is similar to what was obtained for Gaussian stimuli. The mutual information rate I_{noise} also increases with stimulus contrast. However, we have I_{noise} almost equal to I_{spon} for low contrasts and I_{noise} greater than I_{spon} for higher contrasts σ_{SAM}. This occurs for the same reason as described above with Gaussian stimuli: the SNR increases with contrast σ_{SAM} and

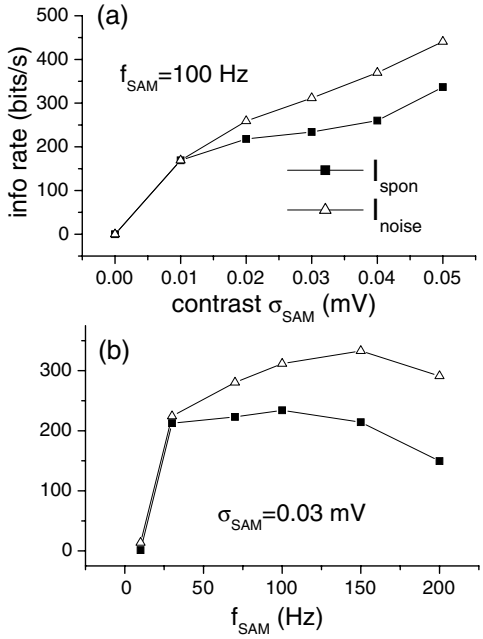

Figure 8. (a) Information measures I_{spon} and I_{noise} as a function of SAM contrast for a constant SAM frequency of 100 Hz. I_{noise} is greater than I_{spon} for high contrasts (see text for explanation). (b) Information measures I_{spon} and I_{noise} as a function of SAM frequency for a constant SAM contrast of 0.03 mV. Note that both information rates exhibit a maximum as a function of SAM frequency.

the noise entropy rate in the presence of background noise $H_{\mathrm{noise}}(R/S)$ decreases towards the noise entropy rate obtained in the absence of background noise $H(R/S)$.

Figure 8(b) shows the measures I_{spon} and I_{noise} obtained by varying the SAM frequency f_{SAM} and keeping the SAM contrast σ_{SAM} constant. We note that the mutual information rate in the presence of background noise I_{noise} and in the absence of background noise I_{spon} are almost equal for low SAM frequencies f_{SAM} but that the difference I_{noise}–I_{spon} increases with f_{SAM}. This is due to the high-pass filtering properties of the electroreceptors reproduced by our model. Moreover, both measures exhibit a maximum as a function of f_{SAM}. This can be explained as follows: neurons typically exhibit low-pass filtering characteristics due to their membrane time constant. However, electroreceptor neurons also exhibit high-pass filtering characteristics. Thus, it might be expected that the neurons exhibit band-pass filtering characteristics. A resonance in the mutual information rate as a function of stimulus frequency f_{SAM} can then be expected (Hutcheon and Yarom 2000). Note however that this resonance was not seen for Gaussian stimuli. This is partly because one merely adds higher frequency components to a Gaussian signal by increasing its cutoff frequency f_{c}. On the other hand, increasing the SAM frequency f_{SAM} shifts the frequency of the unique Fourier component of the signal, thereby probing the neural response to different frequencies. Figure 8(b) can thus be thought of as an 'information tuning curve' for the neuron. Note that this is different from looking at the variation in information transfer of Gaussian stimuli by varying the EOD (i.e. carrier) frequency itself (Longtin and St-Hilaire 2000). This provides another example that the neuronal response to naturalistic stimuli can be very different to the response to low-passed filtered Gaussian white noise stimuli (Rieke *et al* 1995, Lewen *et al* 2001).

4. Discussion

4.1. Summary

In this study, we have compared different information measures for different stimulus ensemble subsets as well as different contrasts, frequencies and bandwidths in the case where the neuron is spontaneously active. Our results show that different information measures can give us quantitatively and in some cases qualitatively different results. Our results are robust to changes in the binwidth used to compute them as well as different extrapolation procedures. However, results obtained with a particular measure must be interpreted carefully. We now discuss all four measures and their implications.

4.2. Comparison of I_{strong} and I_{gaus}

The mutual information rate calculated from a single member S of the stimulus ensemble I_{strong} has been used to quantify the information transfer by neurons in the visual system (Strong *et al* 1998, Reinagel and Reid 2000, Fairhall *et al* 2001, Lewen *et al* 2001, Nirenberg *et al* 2001). We have shown that this generally agrees well with the measure I_{gaus} obtained by averaging over a subset of the stimulus ensemble when the subset W' was the set of zero-mean Gaussian stimuli of contrast σ and cutoff frequency f_{c}. However, some of these studies (Lewen *et al* 2001, Nirenberg *et al* 2001) used naturalistic stimuli and it remains to be shown whether the measure I_{strong} would agree with the mutual information rate computed over the entire stimulus ensemble W.

A mutual information rate estimate about a stimulus S is only valid when the stimulus ensemble W is well sampled (Borst and Theunissen 1999). It remains to be shown whether the subset of all Gaussian stimuli W_{gaus} studied here provides a good sample of the total stimulus ensemble W and if so, in which cases. Thus, it is unclear whether the mutual information rate I_{gaus} calculated over the subset W_{gaus} would agree with the mutual information rate I calculated over the stimulus ensemble W. There is evidence that this might not be the case in sensory neurons as they must be most efficient at coding natural stimuli relevant for the animal's survival. Indeed, natural stimuli have been shown to elicit higher information rates than low-pass filtered Gaussian white noise (Rieke *et al* 1995, Attias and Schreiner 1998).

4.3. Comparison of I_{gaus} and I_{spon}

The total stimulus ensemble W is in general multi-dimensional and is not easy to characterize like W_{gaus} or W_{SAM}. It is therefore useful to have information measures like I_{spon} and I_{noise} that do not depend on the stimulus ensemble used, but rather depend on the properties of the neurons under study and on the particular stimulus used only through the noise entropy rate $H(R/S)$. We have compared the mutual information rate calculated from the spontaneous activity of the neuron I_{spon}, which is identical to the Kullback entropy (Kullback 1959) between the response in the absence of stimulus and the one with a stimulus, to the mutual information rate I_{gaus} calculated over the Gaussian stimulus ensemble W_{gaus} and showed that the two had the same dependence on stimulus contrast and cutoff frequency. The mutual information rate calculated from spontaneous activity I_{spon} was shown to be lower than I_{gaus}. This is because the measure I_{spon} only measures the reduction in the entropy rate of the spike train from spontaneous activity $H_{\mathrm{spon}}(R)$ brought about by a stimulus. This difference is perhaps most clearly seen when no stimulus is present. We then have that $I_{\mathrm{spon}} = 0$ and $I_{\mathrm{gaus}} = H_{\mathrm{gaus}}(R) - H_{\mathrm{spon}}(R) > 0$. As mentioned earlier, the quantity $H_{\mathrm{gaus}}(R) - H_{\mathrm{spon}}(R)$ represents the information transmitted by the neuron about the fact that no stimulus is present in the environment since the 'no-stimulus'

case is just another 'stimulus' in the stimulus ensemble W_{gaus}. The two information rates I_{gaus} and I_{spon} were shown to differ by the quantity $H_{\text{gaus}}(R) - H_{\text{spon}}(R)$ which has no dependence on the particular stimulus S used (it only depends of W_{gaus}). Furthermore, we note that we have that $H(R) - H(R/S) = H(R) - H_{\text{spon}}(R) + H_{\text{spon}}(R) - H(R/S)$. Thus, we have that $I = H(R) - H_{\text{spon}}(R) + I_{\text{spon}}$ using the definitions of I and I_{spon}. Thus, the information rate calculated from the spontaneous activity of the neuron I_{spon} differs by only a constant from the mutual information rate calculated over the full stimulus ensemble I. The two measures will thus behave in the same manner. The measure I_{spon} is zero when there is no stimulus and has the advantage of having no dependence on the stimulus ensemble.

4.4. Conditions of applicability for I_{spon}

It has been argued that a spontaneously active neuron must change one or more properties of its spike train in order to signal the presence of a stimulus (Ratnam and Nelson 2000, Chacron *et al* 2001b). Thus, a change from spontaneous activity might signal the presence of a stimulus for these neurons. One possible way to look at this change is to calculate the difference between the entropy rates of the spontaneous activity and the activity in the presence of a stimulus: this is exactly the information rate I_{spon} which is also known as the Kullback entropy between the neural response in the absence of stimulus and the one with stimulus (Kullback 1959). Such a measure might be applicable to a situation for example in which an animal must detect a stimulus from the absence of stimuli: this occurs for example when the electric fish must detect a prey in midstream (Nelson and MacIver 1999) or when an animal detects a sound in an otherwise silent environment using a spontaneously active auditory afferent.

We now discuss the conditions of applicability of the measure I_{spon}. Clearly, I_{spon} cannot be applied to a neuron that is not spontaneously active. Moreover, it cannot be applied to a neuron with a very regular spike train (e.g. a stretch receptor). There thus needs to be some irregularity in the spontaneous activity of the neuron. However, most spontaneously active neurons have irregular patterns of activity and the trial-to-trial variability decreases when driven by a stimulus.

4.5. Effects of background noise

In nature, stimuli are often contaminated by unwanted noise. Thus, we have introduced a simple extension of the measure I_{spon} that accounts for background noise: I_{noise}. This measure could be applied in the context of the aforementioned cocktail party example. Our results show that, when the stimulus contrast was low, I_{noise} and I_{spon} were approximately equal. However, as stimulus contrast increased, I_{noise} became higher than I_{spon}. This was because the spike train entropy rate in the presence of background noise $H_{\text{noise}}(R)$ was higher than $H_{\text{spon}}(R)$ and because the noise entropy rate in the presence of background noise $H_{\text{noise}}(R/S)$ tended towards $H(R/S)$ as the signal-to-noise ratio increased. This result might seem surprising at first. However, we are only calculating the effect of background noise on the spontaneous activity of a neuron. It has been shown experimentally that an animal could take advantage of background noise for neural coding prey detection (Levin and Miller 1996, Greenwood *et al* 2000). The measure I_{noise} might underlie the means of neural computation used for such a task. In the electrosensory system, I_{noise} might be the appropriate measure for the problem the fish faces when detecting prey in background clutter (e.g. root masses). In the auditory system, this measure might be appropriate for quantifying the information transmitted by a spontaneously active neuron about a stimulus in the presence of background noise. However, we note that this increase might simply be a property of the information measure I_{spon} or of the model we use.

It is important to note that we only considered the case where the background noise was additive (i.e. the noise was added to the stimulus) and independent of the stimulus. This might not be the case in every situation (such as noise caused by another fish's EOD for example). Furthermore, we assumed that the background noise had a flat power spectrum up to 50 Hz and had a contrast of 0.01 mV. This will certainly not be true in general. The strength and frequency content of the background noise will most certainly vary depending on the particular situation the animal is in.

It is known that there are internal noise sources such as channel and synaptic noise that contribute to the variability seen in experimental data. Stochastic processes account for these and other possible noise sources in our model. Such noise sources are responsible for the irregularity of the spontaneous activity of our model neuron and are present in all our simulations. As mentioned earlier, some neurons have no or little spontaneous activity. While I_{spon} could not be applied to such neurons, I_{noise} could in principle be applied if the background noise is sufficiently high so that it might elicit a sufficiently irregular response from that neuron. The background noise would in this case act as internal noise sources.

4.6. Comparison between Gaussian and natural stimuli

Finally, we discuss the differences between natural and Gaussian stimuli. Unrepeated presentations of the latter will always elicit a highly variable neural response due to their high entropy since the Gaussian distribution is the distribution with the highest entropy when the variance of the distribution is constrained (Cover and Thomas 1991, Rieke *et al* 1996). However, natural stimuli seldom obey Gaussian statistics as mentioned earlier. For example, the SAMs used in our study are not Gaussian. These stimuli have zero entropy as they are deterministic in nature and contain only a single Fourier component. Thus, it is not practical to estimate the mutual information rate with I_{strong} since the entropy rate of the spike train when stimulated by an unrepeated SAM is actually equal to the noise entropy rate obtained with that same SAM (i.e. $I_{\mathrm{strong}} = 0$). The stimulus presented must thus have some irregularity in order to apply the measure I_{strong}. The ensemble of SAMs is characterized by the SAM contrast σ_{SAM} and SAM frequency f_{SAM} as well as the probability of occurrence of each SAM. Although it is much easier to work with the SAM ensemble than the full stimulus ensemble W, the characterizing of the SAM ensemble would require much work as the probability of occurrence of each SAM in nature would have to be determined experimentally. However, the measures I_{N} and I_{spon} can easily be calculated as they do not depend on the stimulus ensemble.

4.7. An information resonance

By comparing the measures I_{spon} and I_{noise} obtained for both low-passed filtered Gaussian stimuli and the SAMs, we found that the model neuron exhibited a resonance in information rate when the SAM frequency varied. This resonance can be expected from the combination of high-pass and low-pass filtering characteristics of the electroreceptor neurons (Hutcheon and Yarom 2000). It was however only revealed once we used naturalistic stimuli rather than low-passed filtered Gaussian noise. The reason that low-passed Gaussian noise did not elicit this resonance is that we are merely adding higher frequency components to the signal as we increase the cutoff frequency f_{c}. Note however that the resonance might be revealed using higher cutoff frequencies than the ones used in this study. The responses of a neuron to natural stimuli can be very different from those to artificial stimuli (Rieke *et al* 1995, Attias and Schreiner 1998).

5. Conclusion

In conclusion, we have shown that different information measures are designed to answer different questions and thus give different results. Hence, one must thus carefully consider the measure one uses based on the physiological context. Although the direct method of calculating mutual information rate makes no assumptions about the neural code (Borst and Theunissen 1999), the mutual information rate can only be calculated in practice for very restricted subsets of total stimulus ensemble W. Assumptions must thus be made either in terms of the stimulus ensemble itself or in terms of calculating the information based on the no-stimulus ensemble alone. For neurons that exhibit irregular spontaneous activity, I_{spon} provides a natural information measure that does not require a complicated stimulus ensemble. Further, I_{spon} differs from the Shannon mutual information rate by only a constant: the two information measures thus behave similarly. The mutual information rate in the presence of background noise I_{noise} could also be applied to neurons that are not spontaneously active.

Acknowledgments

We thank J Bastian, A M Oswald and J Benda for useful discussions and careful reading of the manuscript. This research was supported by NSERC (MJC and AL) and CIHR (AL and LM) Canada.

References

Aizenman C D and Linden D J 1999 Regulation of the rebound depolarization and spontaneous firing patterns of deep nuclear neurons in slices of rat cerebellum *J. Neurophysiol.* **82** 1697–709

Attias H and Schreiner C E 1998 Coding of naturalistic stimuli by auditory midbrain neurons *Advances in Neural Information Processing Systems* ed M Jordan, M Kearns and S Solla (Cambridge, MA: MIT Press) pp 103–9

Bastian J 1981 Electrolocation I. How the electroreceptors of *Apteronotus albifrons* code for moving objects and other electrical stimuli *J. Comput. Physiol.* A **144** 465–79

Bender J A, Dimitrov A G and Miller J P 2001 Biophysical constraints on the precision of neural coding *Soc. Neurosci. Abstr.* **27**

Berman N J and Maler L 1998 a Inhibition evoked from primary afferents in the electrosensory lateral line lobe of the weakly electric fish (*Apteronotus leptorhynchus*) *J. Neurophysiol.* **80** 3173–96

Borst A and Theunissen F 1999 Information theory and neural coding *Nature Neurosci.* **2** 947–57

Buracas G T, Zador A M, DeWeese M R and Albright T D 1998 Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex *Neuron* **20** 959–69

Chacron M J 2003 Dynamics leading to optimized information transfer *PhD Thesis* University of Ottawa, Ottawa

Chacron M J, Longtin A and Maler L 2001a Simple models of bursting and non-bursting electroreceptors *Neurocomputing* **38** 129–39

Chacron M J, Longtin A and Maler L 2001b Negative interspike interval correlations increase the neuronal capacity for encoding time-varying stimuli *J. Neurosci.* **21** 5328–43

Chacron M J, Longtin A, St-Hilaire M and Maler L 2000 Suprathreshold stochastic firing dynamics with memory in P-type electroreceptors *Phys. Rev. Lett.* **85** 1576–9

Cherry E C 1953 Some experiments on the recognition of speech, with one and two ears *J. Acoust. Soc. Am.* **25** 975–9

Clague H, Theunissen F and Miller J P 1997 Effects of adaptation on neural coding by primary sensory interneurons in the cricket cercal system *J. Neurophysiol.* **77** 207–20

Cover T and Thomas J 1991 *Elements of Information Theory* (New York: Wiley)

Crampton W G R 1998 Electric signal design and habitat preferences in a species rich assembly of gymnotiform fishes from the upper Amazon basis *Ann. Acad. Bras. Cienc.* **70** 805–47

de Ruyter van Steveninck R R, Lewen G D, Strong S P, Koberle R and Bialek W 1997 Reproducibility and variability in neural spike trains *Science* **275** 1805–8

Dulka J G, Maler L and Ellis W 1995 Androgen-induced changes in electrocommunicatory behaviour are correlated with changes in substance P-like immunoreactivity in the brain of the electric fish *Apteronotus leptorhynchus* *J. Neurosci.* **15** 1879–90

Evarts E V 1964 Temporal patterns of discharge of pyramidal tract neurons during sleep and waking in the monkey *J. Neurophysiol.* **27** 152–71

Fairhall A L, Lewen G D, Bialek W and de Ruyter van Steveninck R R 2001 Efficiency and ambiguity in an adaptive neural code *Nature* **412** 787–92

Gabbiani F 1996 Coding of time varying signals in spike trains of linear and half-wave rectifying neurons *Network Comput. Neural Syst.* **7** 61–85

Gabbiani F and Koch C 1998 Principles of spike train analysis *Methods in Neuronal Modeling: From Ions to Networks* ed C Koch and I Segev (Cambridge, MA: MIT Press) pp 313–60

Gabbiani F, Metzner W, Wessel R and Koch C 1996 From stimulus encoding to feature extraction in weakly electric fish *Nature* **384** 564–7

Gaspard P and Wang X J 1993 Noise, chaos, and $(\varepsilon - \tau)$-entropy per unit time *Phys. Rep.* **235** 291–343

Goldberg J M and Fernández C 1971 Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey: I. Resting discharge and response to constant angular acceleration *J. Neurophysiol.* **34** 635–60

Goychuk I 2001 Information transfer with rate-modulated poisson processes: a simple model for nonstationary stochastic resonance *Phys. Rev.* E **64** 021909

Goychuk I and Hänggi P 2000 Stochastic resonance in ion channels characterized by information theory *Phys. Rev.* E **61** 4272–80

Greenwood P E, Ward L M, Russell D F, Neiman A and Moss F 2000 Stochastic resonance enhances the electrosensory information available to paddlefish for prey capture *Phys. Rev. Lett.* **84** 4773–6

Heiligenberg W, Keller C H, Metzner W and Kawasaki M 1991 Structure and function of neurons in the complex of the nucleus electrosensorius of the gymnotiform fish *Eigenmannia*: detection and processing of electric signals in social communication *J. Comput. Physiol.* A **169** 151–64

Hopkins C D 1973 Lightning as background noise for communication among electric fish *Nature* **242** 268–70

Hubel D 1959 Single-unit activity in striate cortex of unrestrained cats *J. Physiol. Lond.* **147** 226–38

Hutcheon B and Yarom Y 2000 Resonance, oscillation and the intrinsic frequency preferences of neurons *TINS* **23** 216–22

Jaeger D and Bauer J M 1994 Prolonged responses in rat cerebellar Purkinje cells following activation of the granule cell layer: an intracellular *in vitro* and *in vivo* investigation *Exp. Brain Res.* **100** 200–14

Kiang N Y S 1965 *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (Cambridge, MA: MIT Press)

Köppl C 1997 Frequency tuning and spontaneous activity in the auditory nerve and cochlear nucleus magnocellularis of the barn owl tyto alba *J. Neurophysiol.* **77** 364–77

Kreiman G, Krahe R, Metzner W, Koch C and Gabbiani F 2000 Robustness and variability of neuronal coding by amplitude sensitive afferents in the weakly electric fish eigenmania *J. Neurophysiol.* **84** 189–224

Kullback S 1959 *Information Theory and Statistics* (New York: Wiley)

Laughlin S 1981 A simple coding procedure enhances a neuron's information capacity *Z. Naturf.* **36** 910–2

Levin J and Miller J P 1996 Stochastic resonance enhances neural encoding of broadband stimuli in the cricket cercal sensory system *Nature* **380** 165–8

Lewen G D, Bialek W and de Ruyter van Steveninck R R 2001 Neural coding of naturalistic motion stimuli *Network* **12** 317–29

Longtin A and St-Hilaire M 2000 Encoding carrier amplitude modulations via stochastic phase synchronization *Int. J. Bifurcation Chaos* **10** 1–16

Machens C K, Stemmler M B, Prinz P, Krahe R, Ronacher B and Herz A V 2001 Representation of acoustic communication signals by insect auditory neurons *J. Neurosci.* **21** 3215–27

Mainen Z F and Sejnowski T J 1995 Reliability of spike timing in neocortical neurons *Science* **268** 1503–6

Neiman A, Shulgin B, Anishchenko V, Ebelling W, Schimansky-Geier L and Freund J 1996 Dynamical entropies applied to stochastic resonance *Phys. Rev. Lett.* **76** 4299–302

Neiman A and Russell D F 2002 Synchronization of noise-induced bursts in noncoupled sensory neurons *Phys. Rev. Lett.* **88** 138103

Nelson M E and MacIver M A 1999 Prey capture in the weakly electric fish Apteronotus leptorhynchus: sensory acquisition strategies and electrosensory consequences *J. Exp. Biol.* **202** 1195–203

Nirenberg S, Carcieri S M, Jacobs A L and Latham P E 2001 Retinal ganglion cells act largely as independent encoders *Nature* **411** 698–701

Parthridge B L and Heiligenberg W 1980 Three's a crowd? Predicting Eigenmania's response to multiple jamming *J. Comput. Physiol.* **136** 153–64

Ratnam R and Nelson M E 2000 Non-renewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals *J. Neurosci.* **20** 6672–83

Reinagel P and Reid R C 2000 Temporal coding of visual information in the thalamus *J. Neurosci.* **20** 5392–400

Rieke F, Bodnar D A and Bialek W 1995 Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents *Proc. R. Soc.* B **262** 259–65

Rieke F, Warland D, de Ruyter van Steveninck R R and Bialek W 1996 *Spikes: Exploring the Neural Code* (Cambridge, MA: MIT Press)

Roddey J C and Jacobs G A 1996 Information theoretic analysis of dynamical encoding by filiform mechanoreceptors in the cricket cercal system *J. Neurophysiol.* **75** 1365–76

Shannon C E 1948a The mathematical theory of communication *Bell Syst. Tech. J.* **27** 379–423

Shannon C E 1948b The mathematical theory of communication *Bell Syst. Tech. J.* **27** 623–56

Simoncelli E P and Olshausen B A 2001 Natural image statistics and neural representation *Annu. Rev. Neurosci.* **24** 1193–216

Steriade M 1978 Cortical long-axoned cells and putative interneurons during the sleep-waking cycle *Behav. Brain Res.* **3** 465–514

Steuer R, Ebeling W, Russell D F, Bahar S, Neiman A and Moss F 2001a Entropy and local uncertainty of data from sensory neurons *Phys. Rev.* E **64** 0619111–6

Steuer R, Molgedey L, Ebelling W and Jiménez-Montaño M A 2001b Entropy and optimal partition for data analysis *Eur. Phys. J.* B **19** 265–9

Strong S P, Koberle R, de Ruyter van Steveninck R R and Bialek W 1998 Entropy and information in neural spike trains *Phys. Rev. Lett.* **80** 197–200

Tiesinga P H E 2001 Information transmission and recovery in neural communication channels revisited *Phys. Rev.* E **64** 012901

Warland D K, Reinagel P and Meister M 1997 Decoding visual information from a population of retinal ganglion cells *J. Neurophysiol.* **78** 2336–50

Wessel R, Koch C and Gabbiani F 1996 Coding of time-varying electric field amplitude modulations in a wave-type electric fish *J. Neurophysiol.* **75** 2280–93

Xu Z, Payne J R and Nelson M E 1996 Logarithmic time course of sensory adaptation in electrosensory afferent nerve fibers in a weakly electric fish *J. Neurophysiol.* **76** 2020–32

Zador A 1998 Impact of synaptic unreliability on the information transmitted by spiking neurons *J. Neurophysiol.* **79** 1219–29

Zakon H H 1986 The electroreceptive periphery *Electroreception* ed T H Bullock and W Heiligenberg (New York: Wiley) pp 103–56

Zupanc G K H and Maler L 1993 Evoked chirping in the weakly electric fish *Apteronotus leptorhynchus*: a quantitative biophysical analysis *Can. J. Zoo.* **71** 2301–10