

## Entropy and Information in Neural Spike Trains

S. P. Strong,<sup>1</sup> Roland Koberle,<sup>1,2</sup> Rob R. de Ruyter van Steveninck,<sup>1</sup> and William Bialek<sup>1</sup>

<sup>1</sup>*NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540*

<sup>2</sup>*Department of Physics, Princeton University, Princeton, New Jersey 08544*

(Received 29 February 1996; revised manuscript received 11 July 1997)

The nervous system represents time dependent signals in sequences of discrete, identical action potentials or spikes; information is carried only in the spike arrival times. We show how to quantify this information, in bits, free from any assumptions about which features of the spike train or input signal are most important, and we apply this approach to the analysis of experiments on a motion sensitive neuron in the fly visual system. This neuron transmits information about the visual stimulus at rates of up to 90 bits/s, within a factor of 2 of the physical limit set by the entropy of the spike train itself. [S0031-9007(97)04939-9]

PACS numbers: 87.10.+e, 05.20.-y, 89.70.+c

As you read this text, optical signals reaching your retina are encoded into sequences of identical pulses, termed action potentials or spikes, that propagate along the  $\sim 10^6$  fibers of the optic nerve from the eye to the brain. This encoding appears universal, occurring in animals as diverse as worms and humans, and spanning all the sensory modalities [1]. The molecular mechanisms for the generation and propagation of action potentials are well understood [2], as are the mathematical reasons for the selection of stereotyped pulses [3]. Less well understood is the function of these spikes as a code [4]: How do sequences of spikes represent the sensory world, and how much information is conveyed in this representation?

The temporal sequence of spikes provides a large capacity for transmitting information [5]. One central question is whether the brain takes advantage of this large capacity, or whether variations in spike timing represent noise which must be averaged away [4,6]. In response to a long sample of time varying stimuli, the spike train of a single neuron varies, and we can quantify this variability by the entropy per unit time of the spike train,  $S(\Delta\tau)$  [7], which depends on the time resolution  $\Delta\tau$  with which we record the spike arrival times. If we repeat the same time dependent stimulus, we see a similar, but not precisely identical, sequence of spikes (Fig. 1). This variability at fixed input can also be quantified by an entropy, which we call the conditional or noise entropy per unit time  $\mathcal{N}(\Delta\tau)$ . The information that the spike train provides about the stimulus is the difference between these entropies,  $R_{\text{info}} = S - \mathcal{N}$  [7]. Because  $\mathcal{N}$  is positive (semi)definite,  $S$  sets the capacity for transmitting information, and we can define an efficiency  $\epsilon(\Delta\tau) = R_{\text{info}}(\Delta\tau)/S(\Delta\tau)$  with which this capacity is used [9]. The question of whether spike timing is important is really the question of whether this efficiency is high at small  $\Delta\tau$  [4].

For some neurons, we understand enough about what the spike train represents that direct “decoding” of the spike train is possible; the information extracted by these decoding methods can be more than half of the total spike train entropy with  $\Delta\tau \sim 1$  ms [9]. The idea that sen-

sory neurons provide a maximally efficient representation of the outside world has also been suggested as an optimization principle from which many features of these cells’ responses can be derived [10]. But, particularly in the central nervous system [6], assumptions about what is being encoded should be viewed with caution. The goal of this paper is, therefore, to give a completely model independent estimate of entropy and information in neural spike trains as they encode dynamic signals.

We begin by discretizing the spike train into time bins of size  $\Delta\tau$ , and examining segments of the spike train in windows of length  $T$ , so that each possible neural response is a “word” with  $T/\Delta\tau$  symbols. Let us call the normalized count of the  $i$ th word  $\tilde{p}_i$ . Then the “naive estimate” of the entropy is

$$S_{\text{naive}}(T, \Delta\tau; \text{size}) = - \sum_i \tilde{p}_i \log_2 \tilde{p}_i; \quad (1)$$

the notation reminds us that our estimate depends on the size of the data set. The true entropy is

$$S(T, \Delta\tau) = \lim_{\text{size} \rightarrow \infty} S_{\text{naive}}(T, \Delta\tau; \text{size}), \quad (2)$$

and we are interested in the entropy rate  $S(\Delta\tau) = \lim_{T \rightarrow \infty} S(T, \Delta\tau)/T$ . At large  $T$ , very large data sets are required to ensure convergence of  $S_{\text{naive}}$  to the true entropy  $S$ . Imagine a spike train with mean spike rate  $\bar{r} \sim 40$  spikes/s, sampled with a time resolution  $\Delta\tau = 3$  ms. In a window of  $T = 100$  ms, the maximum entropy consistent with this mean rate [4,5] is  $S \sim 17.8$  bits, and the entropy of real spike trains is not far from this bound. But then there are roughly  $2^S \sim 2 \times 10^5$  words with significant  $p_i$ , and our naive estimation procedure seems to require that we observe many samples of each word. If our samples come from nonoverlapping 100 ms windows, then we need much more than 3 hours of data.

It is possible to make progress despite these pessimistic estimates. First, we examine explicitly the dependence of  $S_{\text{naive}}$  on the size of the data set and find regular behaviors [11] that can be extrapolated to the infinite data limit. Second, we evaluate robust bounds [7,12] on the entropy that serve as a check on our extrapolation procedure.

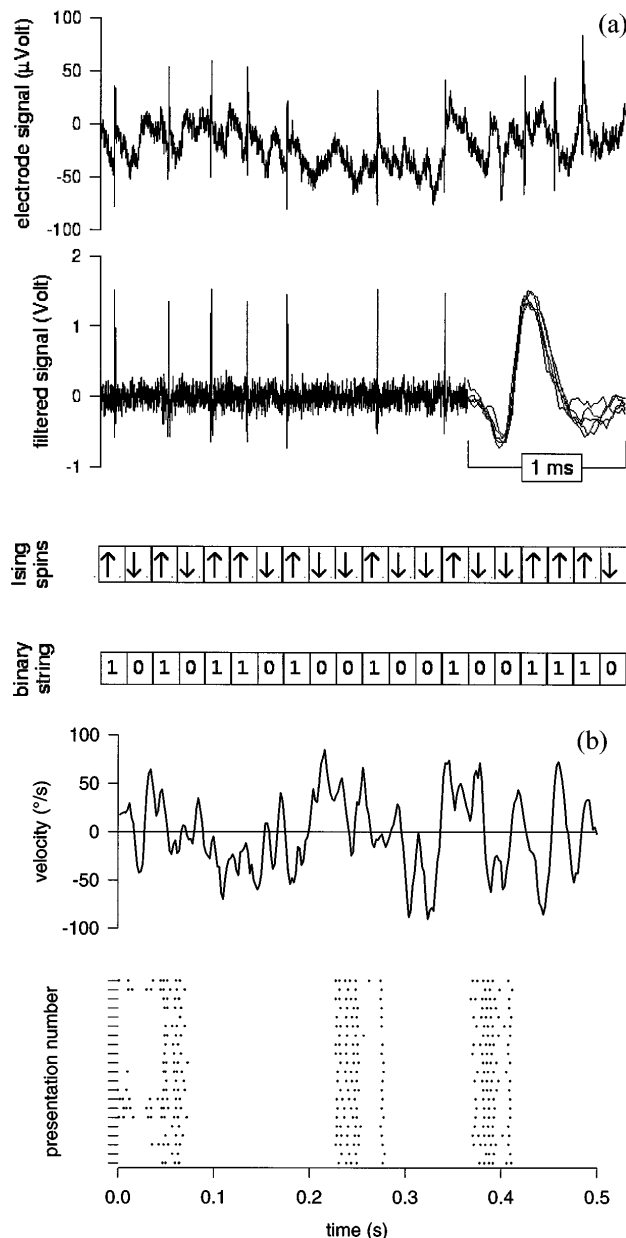


FIG. 1. (a) Raw voltage records from a tungsten microelectrode near the cell H1 are filtered to isolate the action potentials. The expanded scale shows a superposition of several spikes to illustrate their stereotyped form. (b) Angular velocity of a pattern moving across the fly's visual field produces a sequence of spikes in H1, indicated by dots. Repeated presentations produce slightly different spike sequences. For experimental methods, see Ref. [8].

Third, we are interested in the extensive component of the entropy, and we find that a clean approach to extensivity is visible before sampling problems set in. Finally, for the neuron studied—the motion sensitive neuron H1 in the fly's visual system—we can actually collect many hours of data.

H1 responds to motion across the entire visual field, producing more spikes for an inward horizontal motion and fewer spikes for an outward motion; vertical motions are coded by other neurons [13]. These cells provide vi-

sual feedback for flight control. In the experiments analyzed here, the fly is immobilized and views computer generated images on a display oscilloscope. For simplicity these images consist of vertical stripes with randomly chosen grey levels, and this pattern takes a random walk in the horizontal direction [14].

We begin our analysis with time bins of size  $\Delta\tau = 3$  ms. For a window of  $T = 30$  ms—corresponding to the behavioral response time of the fly [15]—Fig. 2 shows the histogram  $\{\tilde{p}_i\}$ , and the naive entropy estimates. We see that there are very small finite data set corrections ( $<10^{-3}$ ), well fit by [11]

$$S_{\text{naive}}(T, \Delta\tau; \text{size}) = S(T, \Delta\tau) + \frac{S_1(T, \Delta\tau)}{\text{size}} + \frac{S_2(T, \Delta\tau)}{\text{size}^2}. \quad (3)$$

Under these conditions we feel confident that the extrapolated  $S(T, \Delta\tau)$  is the correct entropy. For sufficiently large  $T$ , finite size corrections are larger, the contribution of the second correction is significant, and the extrapolation to infinite size is unreliable.

Ma [12] discussed the problem of entropy estimation in the undersampled limit. For probability distributions that are uniform on a set of  $N$  bins (as in the microcanonical ensemble), the entropy is  $\log_2 N$  and the problem is to estimate  $N$ . Ma noted that this could be done by counting

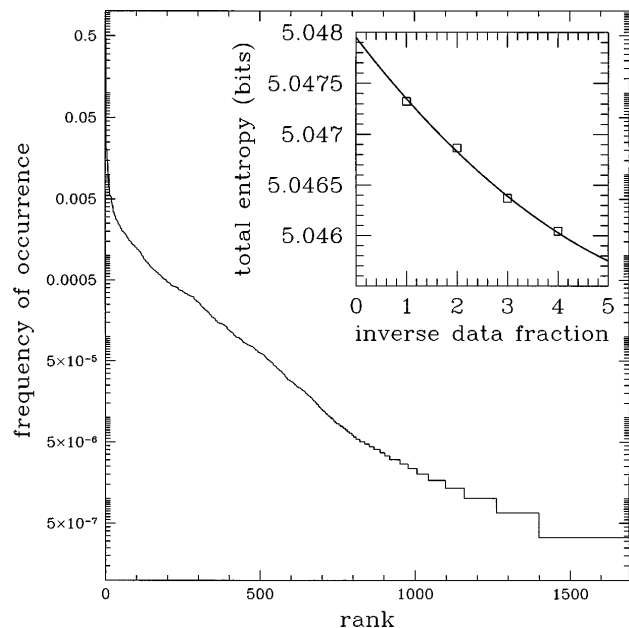


FIG. 2. The frequency of occurrence for different words in the spike train, with  $\Delta\tau = 3$  ms and  $T = 30$  ms. Words are placed in order so that the histogram is monotonically decreasing; at this value of  $T$  the most likely word corresponds to no spikes. Inset shows the dependence of the entropy, computed from this histogram according to Eq. (1), on the fraction of data included in the analysis. Also plotted is a least squares fit to the form  $S = S_0 + S_1/\text{size} + S_2/\text{size}^2$ . The intercept  $S_0$  is our extrapolation to the true value of the entropy with infinite data [11].

the number of times that two randomly chosen observations yield the same configuration, since the probability of such a coincidence is  $1/N$ . More generally, the probability of a coincidence is  $P_c = \sum p_i^2$ , and hence

$$S = - \sum p_i \log_2 p_i = - \langle \log_2 p_i \rangle, \\ \geq - \log_2 \langle p_i \rangle = - \log_2 P_c, \quad (4)$$

so we can compute a lower bound to the entropy by counting coincidences. Furthermore,  $\log_2 P_c$  is less sensitive to sampling errors than is  $S_{\text{naive}}$ . The Ma bound is one of the Renyi entropies [16] widely used in the analysis of dynamical systems [17].

The Ma bound is tightest for distributions that are close to uniform, but distributions of neural responses cannot be uniform because spikes are sparse. The distribution of words with fixed spike count is more nearly uniform, so we use  $S \geq S_{\text{Ma}}$ , with

$$S_{\text{Ma}} = - \sum_{N_{\text{sp}}} P(N_{\text{sp}}) \\ \times \log_2 \left[ P(N_{\text{sp}}) \frac{2n_c(N_{\text{sp}})}{N_{\text{obs}}(N_{\text{sp}})[N_{\text{obs}}(N_{\text{sp}}) - 1]} \right], \quad (5)$$

where  $n_c(N_{\text{sp}})$  is the number of coincidences observed among the words with  $N_{\text{sp}}$  spikes,  $N_{\text{obs}}(N_{\text{sp}})$  is the total number of occurrences of words with  $N_{\text{sp}}$  spikes, and  $P(N_{\text{sp}})$  is the fraction of words with  $N_{\text{sp}}$  spikes.

In Fig. 3 we plot our entropy estimate as a function of the window size  $T$ . For sufficiently large windows the naive procedure gives an answer smaller than the Ma bound, and hence the naive answer must be unreliable because it is more sensitive to sampling problems. For smaller windows the lower bound and the naive estimate are never more than 10%–15% apart. The point at which the naive estimate crashes into the Ma bound is also where the second correction in Eq. (3) becomes significant and we lose control over the extrapolation to the infinite data limit,  $T \sim 100$  ms. Beyond this point the Ma bound becomes steadily less powerful.

If the correlations in the spike train have finite range, then the leading subextensive contribution to the entropy will be a constant  $C(\Delta\tau)$ ,

$$\frac{S(T, \Delta\tau)}{T} = S(\Delta\tau) + \frac{C(\Delta\tau)}{T} + \dots \quad (6)$$

This behavior is seen clearly in Fig. 3, and emerges before the sampling disaster, providing an estimate of the entropy per unit time,  $S(\Delta\tau = 3 \text{ ms}) = 157 \pm 3$  bits/s.

The entropy approaches its extensive limit from above [7], so that

$$\frac{1}{\Delta\tau} [S(T + \Delta\tau, \Delta\tau) - S(T, \Delta\tau)] \geq S(\Delta\tau) \quad (7)$$

for all window sizes  $T$ . This bound becomes progressively tighter at larger  $T$ , until sampling problems set in. In fact there is a broad plateau ( $\pm 2.7\%$ ) in the range  $18 < T < 60$  ms, leading to  $S \leq 157 \pm 4$  bits/s, in excellent agreement with the extrapolation in Fig. 3.

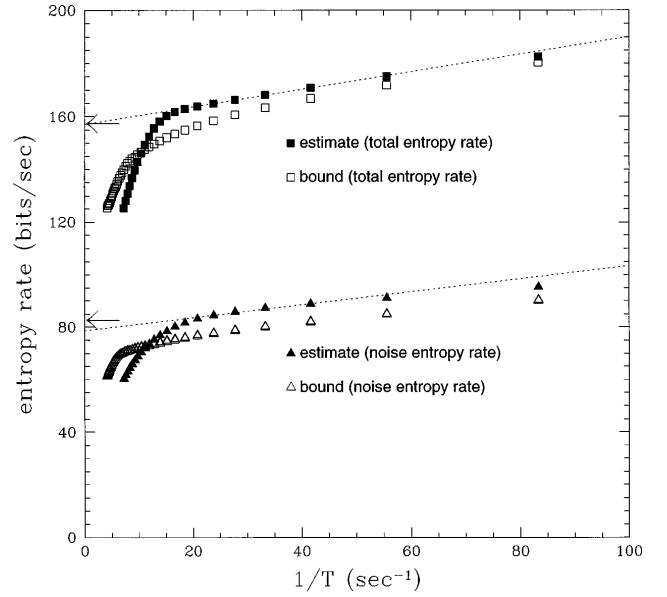


FIG. 3. The total and noise entropies per unit time are plotted versus the reciprocal of the window size, with the time resolution held fixed at  $\Delta\tau = 3$  ms. Results are given both for the direct estimate and for the bounding procedure described in the text, and for each data point we apply the extrapolation procedures of Fig. 2 (inset). Dashed lines indicate extrapolations to infinite word length, as discussed in the text, and arrows indicate upper bounds obtained by differentiating  $S(T)$  [7].

The noise entropy per unit time  $\mathcal{N}(\Delta\tau)$  measures the variability of the spike train in response to repeated presentations of the same time varying input signal. Marking a particular time  $t$  relative to the stimulus, we accumulate the frequencies of occurrence  $\tilde{p}_i(t)$  of each word  $i$  that begins at  $t$ , then form the naive estimate of the local noise entropy in the window from  $t$  to  $t + T$ ,

$$N_{\text{naive}}^{\text{local}}(t, T, \Delta\tau; \text{size}) = - \sum_i \tilde{p}_i(t) \log_2 \tilde{p}_i(t). \quad (8)$$

The average rate of information transmission by the spike train depends on the noise entropy averaged over  $t$ ,  $N_{\text{naive}}(T, \Delta\tau; \text{size}) = \langle N_{\text{naive}}^{\text{local}}(t; T, \Delta\tau, \text{size}) \rangle_t$ . Then we analyze as before the extrapolation to large data set size and large  $T$ , with the results shown in Fig. 3. The difference between the two entropies is the information which the cell transmits,  $R_{\text{info}}(\Delta\tau = 3 \text{ ms}) = 78 \pm 5$  bits/s, or  $1.8 \pm 0.1$  bits/spike.

This analysis has been carried out over a range of time resolutions,  $800 > \Delta\tau > 2$  ms, as summarized in Fig. 4. Note that  $\Delta\tau = 800$  ms corresponds to counting spikes in bins that contain typically thirty spikes, while  $\Delta\tau = 2$  ms corresponds to timing each spike to within 5% of the typical interspike interval. Over this range, the entropy of the spike train varies over a factor of roughly 40, illustrating the increasing capacity of the system to convey information by making use of spike timing. The information that the spike train conveys about the visual stimulus increases in approximate proportion to

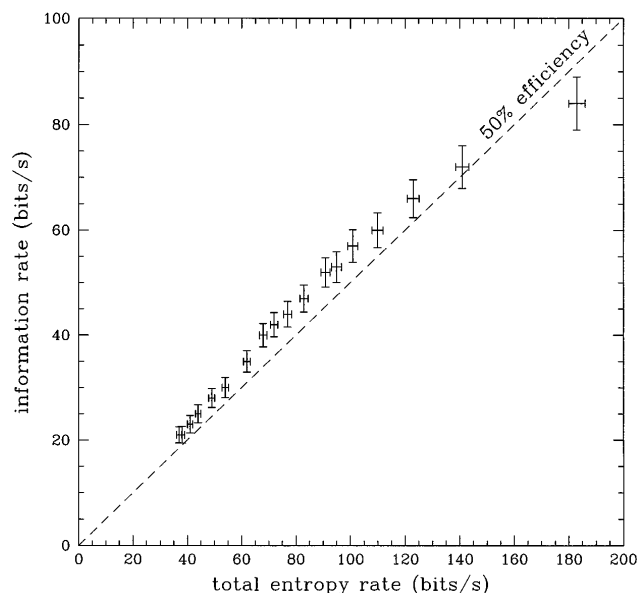


FIG. 4. Information and entropy rates computed at various time resolutions, from  $\Delta\tau = 800$  ms (lower left) to  $\Delta\tau = 2$  ms (upper right). Note the approximately constant, high efficiency across the wide range of entropies. For comparison, the dashed line is the line of 50% efficiency.

the entropy, corresponding to  $\sim 50\%$  efficiency, at least for this ensemble of stimuli.

Although we understand a good deal about the signals represented in H1 [13,18], our present analysis does not hinge on this knowledge. Similarly, although it is possible to collect very large data sets from H1, Figs. 2 and 3 suggest that more limited data sets would compromise our conclusions only slightly. It is feasible, then, to apply these same analysis techniques to cells in the mammalian brain [19]. Like cells in the monkey or cat primary visual cortex, H1 is four layers “back” from the array of photodetectors and receives its input from thousands of synapses. For this central neuron, half the available information capacity is used, down to millisecond precision. Thus, the analysis developed here allows us for the first time to demonstrate directly the significance of spike timing in the nervous system without any hypotheses about the structure of the neural code.

We thank N. Brenner, K. Miller, and P. Mitra for helpful discussions, and G. Lewen for his help with the experiments. R.K., on leave from the Institute of Physics, Universidade de São Paulo, São Carlos, Brasil, was supported in part by CNPq and FAPESP.

- [1] E. D. Adrian, *The Basis of Sensation* (W. W. Norton, New York, 1928).
- [2] D. J. Aidley, *The Physiology of Excitable Cells* (Cambridge University Press, Cambridge, 1989), 3rd ed.
- [3] A. L. Hodgkin and A. F. Huxley, *J. Physiol.* **117**, 500 (1952); D. G. Aronson and H. F. Weinberger, *Adv. Math.* **30**, 33 (1978).

- [4] F. Rieke, D. Warland, R. R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, 1997).
- [5] D. MacKay and W. S. McCulloch, *Bull. Math. Biophys.* **14**, 127 (1952).
- [6] See, for example, the recent commentaries: M. N. Shadlen and W. T. Newsome, *Curr. Opin. Neurobiol.* **5**, 248 (1995); W. R. Softky, *ibid.* **5**, 239 (1995); T. J. Sejnowski, *Nature (London)* **376**, 21 (1995); D. Ferster and N. Spruston, *Science* **270**, 756 (1995); C. F. Stevens and A. Zador, *Curr. Biol.* **5**, 1370 (1995).
- [7] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).
- [8] R. R. de Ruyter van Steveninck and W. Bialek, *Proc. R. Soc. London B* **234**, 379 (1988).
- [9] F. Rieke, D. Warland, and W. Bialek, *Europhys. Lett.* **22**, 151 (1993); F. Rieke, D. A. Bodnar, and W. Bialek, *Proc. R. Soc. London B* **262**, 259 (1995).
- [10] J. J. Atick, in *Princeton Lectures on Biophysics*, edited by W. Bialek (World Scientific, Singapore, 1992), pp. 223–289.
- [11] A. Treves and S. Panzeri, *Neural Comput.* **7**, 399 (1995).
- [12] S. K. Ma, *J. Stat. Phys.* **26**, 221 (1981).
- [13] For reviews, see N. Franceschini, A. Riehle, and A. le Nestour, in *Facets of Vision*, edited by D. G. Stavenga and R. C. Hardie (Springer-Verlag, Berlin, 1989), Chap. 17; K. Hausen and M. Egelhaaf, *ibid.*, Chap. 18. For an overview of the fly’s anatomy, see the contribution by N. J. Strausfeld, *ibid.*, Chap. 16.
- [14] The entropy and information in spike trains obviously depend on stimulus conditions. In these experiments the fly is immobilized 13.24 cm from the screen of a display oscilloscope and views images with an angular area of  $875 \text{ deg}^2$ , which stimulates 3762 photoreceptors. The mean radiance of the screen is  $180 \text{ mW}/(\text{sr m}^2)$ , or  $\sim 5 \times 10^4$  photons/s in each photoreceptor, roughly the light level for a fly flying at dusk. The pattern on the screen consists of bars randomly set dark or light, with the bar width equal to the receptor lattice spacing. Finally, the pattern takes a random walk in 2 ms time steps, corresponding to the refresh time of the monitor; the diffusion constant is  $2.8 \text{ deg}^2/\text{s}$ .
- [15] M. F. Land and T. S. Collett, *J. Comp. Physiol.* **89**, 331 (1974).
- [16] A. Renyi, *Probability Theory* (North-Holland, Amsterdam, 1970).
- [17] D. Ruelle, *Chaotic Evolution and Strange Attractors* (Cambridge University Press, Cambridge, England, 1989).
- [18] W. Bialek, F. Rieke, R. R. de Ruyter van Steveninck, and D. Warland, *Science* **252**, 1757 (1991).
- [19] We have analyzed a limited data set from experiments by K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon, *Visual Neurosci.* **10**, 1157 (1993), who studied the response of motion sensitive cells in the monkey visual cortex; a discussion of these data is given by W. Bair and C. Koch, *Neural Comput.* **8**, 1185 (1996). For the experiment in Fig. 1B of Bair and Koch, we find information rates of  $\sim 2$  bits/spike at  $\Delta\tau = 5$  ms, and this result emerges clearly from just three minutes of data. We thank W. Bair for making these data available to us.