

Energy Efficient Neural Codes

William B Levy
Robert A. Baxter

Department of Neurosurgery, University of Virginia,
Charlottesville, VA 22908 USA

In 1969 Barlow introduced the phrase "economy of impulses" to express the tendency for successive neural systems to use lower and lower levels of cell firings to produce equivalent encodings. From this viewpoint, the ultimate economy of impulses is a neural code of minimal redundancy. The hypothesis motivating our research is that energy expenditures, e.g., the metabolic cost of recovering from an action potential relative to the cost of inactivity, should also be factored into the economy of impulses. In fact, coding schemes with the largest representational capacity are not, in general, optimal when energy expenditures are taken into account. We show that for both binary and analog neurons, increased energy expenditure per neuron implies a decrease in average firing rate if energy efficient information transmission is to be maintained.

1 Introduction

A theory of neuronal encoding (Barlow 1959) hypothesizes that minimizing the representational entropy is highly desirable and, therefore, is a driving force in the evolution of neuronal codes. Clearly many neural computations can be interpreted from this perspective. However, from an engineering perspective, natural selection might have to make some compromises. For example, the brain is one of the metabolically most active organs of the body. In children, when a large number of neuronal codes are developing, the brain accounts for up to 50% of the resting total body oxygen consumption (Sokoloff 1989). Because the brain uses so much energy, there is the problem of energy efficient use of neurons. Thus, neuronal representations may be optimal neither for the information they carry nor for the energy they use; rather there might be some sort of optimal compromise. Here we will examine this combination of concerns.

One of the major goals of neuroscience is to understand the codes used by the brain for making sense of the endless stream of signals encountered as it interacts with the environment. One pivotal design consideration for any information processing system is the *representational*

capacity of the coding scheme, i.e., the number of distinct patterns that can be represented. Ever since Barlow's introduction of the economy of impulses (Barlow 1969), researchers have tended to concentrate on codes that minimize redundant information and maximize representational capacity (Atick 1992; Adelsberger-Mangan and Levy 1992; Foldiak 1990; Redlich 1993).

We (Levy 1985) certainly agree with Barlow that low redundancy codes are desirable, but pure coding considerations (e.g., redundancy or statistical dependence) may not be the sole parameter driving the evolution of neural codes (Barlow 1959). Here, we promote the hypothesis that energy expenditures should also be factored into the economy of impulses.¹ Consideration of the relative metabolic cost of generating an action potential versus inactivity over an equivalent time period leads to a unique maximization of the ratio of representational capacity to energy expended. This maximization thus sets the spike frequency.

We consider this maximization for both binary and analog neurons. The binary case is the simplest case because each neuron has only two possible states: on or off. For this reason, we first discuss the binary case in Section 2 and leave the analog case for Section 3.

2 Case 1: Binary Neurons

Under the constraint that a fraction of the cells are active on average for any given stimulus, we develop an expression for the representational capacity of a neuronal population composed of n neurons with a fraction of the cells active on average for any given stimulus. Next we develop an expression for the energy expended to support such a neural code. Finally, maximization of the ratio of the representational capacity to the energy expended determines how many cells should be active on average (i.e., the firing probabilities).

2.1 Representational Capacity. We first obtain an expression for the representational capacity of a population of neurons with a fraction of the cells firing. This result generalizes to the case in which any number of cells is firing, provided p represents the average fraction of cells firing or the firing probability of any cell in the neuronal population.

In distributed coding schemes with a fixed number of cells active for any given stimulus, a critical question is: How many coding cells should be active relative to the number of cells available? For example, the so called "grandmother" coding scheme allows only one cell to be active for any given stimulus and has a very limited representational capacity—the number of cells available is the number of distinct patterns that can be represented. In contrast, if half of all available cells are active on

¹We thank the reviewers for pointing out that Softky and Kammen (1991) mentioned this hypothesis previously.

average then the representational capacity is maximized, assuming the neurons are binary signaling devices. Thus, for a coding scheme with a fixed number of cells active, the optimum coding strategy (in the sense of capacity) requires half of all available cells to fire for any given stimulus.

To obtain an expression for the representational capacity of a population of neurons with a specified fraction of the cells firing on average, we begin with two simplifying assumptions. In this section, we assume that each cell in a population of neurons acts like a binary signaling device. Second, we assume that the number of neurons is large. The maximum representational capacity of a population of n cells with binary outputs is 2^n codewords, or n bits, when the number of active cells is unrestricted. The representational capacity of a population of n cells with only np cells active over a given time interval τ is $n!/[(np)!(n-np)!]$ codewords, which is simply the expression for the number of all possible codes of length n with exactly np cells active, where p is any number on the closed interval $[0, 1]$ such that np is an integer.

The representational capacity expressed in bits per unit time is obtained by taking the base 2 logarithm of the number of codewords:

$$\begin{aligned} C(n, np) &= \log_2 \left[\frac{n!}{(np)!(n-np)!} \right] \\ &= nH(p) - \log_2(\sigma\sqrt{2\pi}) \approx nH(p) \end{aligned} \quad (2.1)$$

where $H(p)$ is Shannon's entropy function of a binary event with probability p ; $H(p) \equiv -p \log_2 p - (1-p) \log_2 (1-p)$. Here, σ is the standard deviation of the binomial distribution, i.e., $\sigma = \sqrt{np(1-p)}$. The simplification $C(n, np) \approx nH(p)$ is based on two approximations. First, we used Stirling's formula (Mathews and Walker 1970) to compute the factorials, e.g., $n!$ is approximately $\sqrt{2\pi}n^{n+1/2}e^{-n}$. For $n \geq 100$, this approximation yields less than 0.1% error. Second, since n is quite large in neuronal networks, the term $\log_2(\sigma\sqrt{2\pi})$ is dropped because, as n becomes increasingly large, this term increases as $\log_2(n)$ whereas $nH(p)$ grows at the rate n . If the code word length is allowed to vary as a binomial distribution around a mean value, the same result is obtained.

2.2 Energy Expenditure. Next we need an expression for the energy expended per unit time when a code uses an average of np neurons active with the remaining $n(1-p)$ neurons inactive. Assume that a resting neuron expends one energy unit in one time unit τ , and that it expends r units of energy when it fires in a time interval τ . Then the average amount of energy expended per time unit for a code with np cells active on average is

$$E = n(1-p) + npr = n[1 + p(r-1)] \quad (2.2)$$

We have left r as a parameter since the value of r will vary from one neuronal type to another. However, existing data can be used to estimate

a reasonable range of values. Sokoloff's laboratory has estimated energy consumption in the rat by measuring glucose utilization in the superior cervical ganglion as a function of stimulation frequency in peripheral sympathetic nerves (Yarowsky *et al.* 1985; Yarowsky *et al.* 1983) and in the spinal cord (Kadekaro *et al.* 1985). Because of the poor frequency following characteristic of the sympathetic nervous system, we have used only the low frequency portion of their curves to estimate the value of r . In the case of the carotid nerve, the value of r is approximately 40; for the spinal cord, r is much higher, approximately 160. Earlier experiments by Ritchie and Straub (1980) in which oxygen consumption and phosphate efflux were measured in sensory nerve fibers yield estimates of r as approximately 30 and 75, depending upon which of the two measures is used. Therefore, it seems reasonable to hypothesize that values of r range from 10 to 200.

2.3 Maximizing C/E . We now form the ratio of representational capacity to energy expended, expressed in bits/energy unit, as

$$\frac{C}{E} \approx \frac{H(p)}{1 + p(r-1)} \quad (2.3)$$

and determine what values of p maximize this ratio. We refer to networks that maximize this ratio as energy efficient. Note that the time interval assumed for both C and E cancels, and, because n also cancels, that the approximation for this ratio is independent of n . The lack of dependence on n implies that *the maximization of C/E can occur locally* (i.e., on a neuron by neuron basis).

To see that this ratio makes sense, consider the improbable case in which the metabolic cost of generating an action potential is the same as the cost of staying at rest, i.e., $r = 1$. For this case, the maximum of C/E occurs at $p = 0.5$, i.e., when only half of the neurons are active on average. This is the value of p that maximizes the representational capacity exclusively. As r increases, however, the optimal value of p decreases. The value of C/E is graphed as a function of p for $r = 1, 10$, and 100 in Figure 1. The maximum C/E occurs at $p = 0.16$ for $r = 10$, $p = 0.03$ for $r = 100$, and $p = 0.005$ for $r = 1000$. Figure 2 shows how the value of p that maximizes C/E decreases as a function of r . Since our calculation from the energy consumption data implies that r ranges from 10 to 200, the value of p that maximizes C/E should range from 0.16 to 0.02. For many cells in the brain, 400 Hz can be used as a reasonable estimate of the maximum firing frequency. Then, with an assumed maximal frequency of 400 Hz, the implied average activity levels range from 64 down to 8 Hz. We note that this range of activity levels is quite close to experimental observations (Sharp and Green 1994).

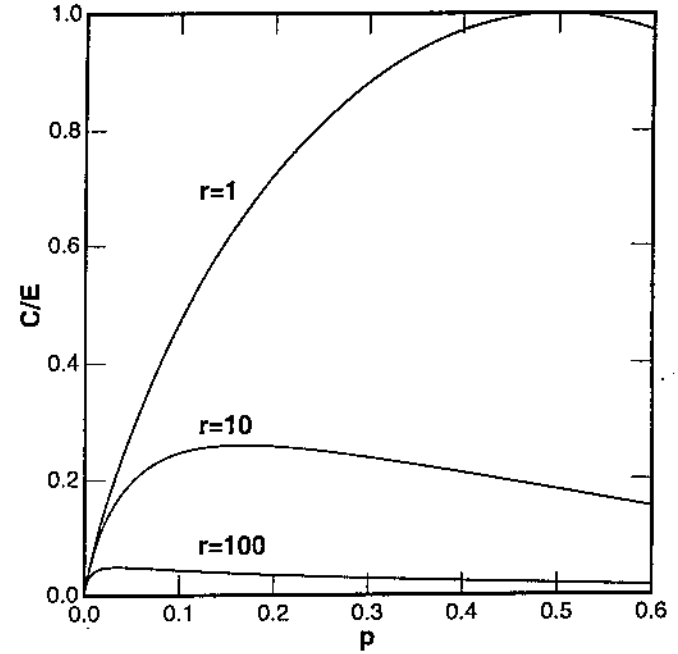


Figure 1: The ratio of representational capacity to energy expenditure (C/E) is a concave function of the average firing probability (p). Here we plot three such curves corresponding to three different values of r , the relative energy expended for an action potential compared to remaining at rest. Note how the maxima across curves decrease as action potentials become more energetically expensive.

3 Case 2: Analog Neurons

In the binary case, we assumed there were n cells with np cells active. We then obtained expressions for the representational (information) capacity of the n neurons and their energy expenditure. From this result, we could simply generalize the notion of "active cells" to cells with firing frequencies greater than some threshold frequency (perhaps corresponding to the spontaneous firing frequency). However, this extension is trivial since the code remains binary. A much more interesting and useful extension is to consider neurons with firing frequencies that span a known range. The firing frequencies of such analog neurons are thought of as

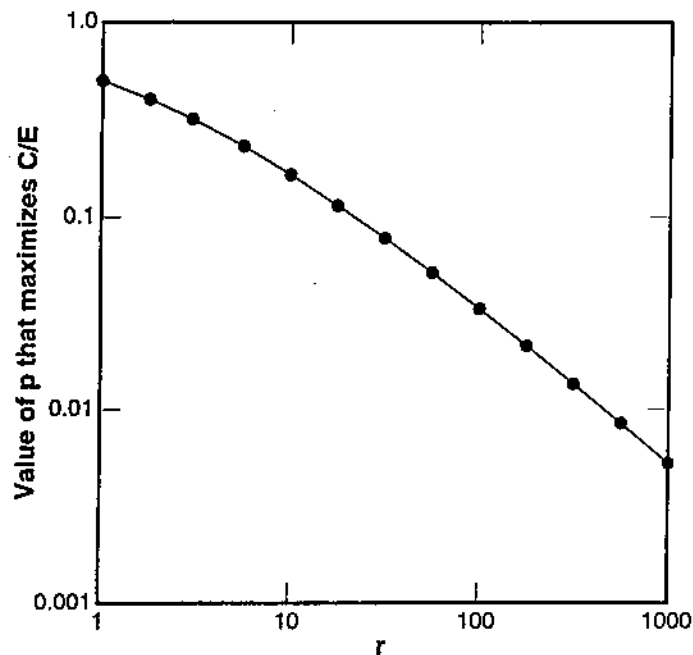


Figure 2: As the relative energy expended for an action potential compared to remaining at rest (r) increases, the optimal firing probabilities for maximizing the ratio of representational capacity to energy expended (C/E) decrease. This decrease is approximately linear on logarithmic scales and can be approximated as $p_{\max} = 0.5r^{-0.7}$. The filled circles in this graph represent maxima of C/E for different values of r . These maxima were determined using the bisection method to solve for the roots of $\partial(C/E)/\partial p = 0$.

their activity levels. Therefore, in the case of analog neurons, we seek to maximize the efficiency given the mean firing frequency.

3.1 Representational Capacity. We start by specifying a time interval T over which the spike frequency is determined. Each of the n cells has a firing frequency f_j determined by j/T , where j is the spike count over the time interval T . The representational capacity is infinite if any number of spikes can be generated within the interval T —regardless of the value of T provided T is nonzero. However, for biological neurons we know that j has an upper bound N —which results in an upper bound on the f_j 's,

say f_N . Therefore, we must determine the representational capacity of n neurons, each of which can take on only the discrete firing frequencies $f_j = \{0, 1/T, 2/T, \dots, N/T\}$. Let the probability that a neuron emits j spikes in the interval T be denoted by p_j . Then the representational capacity per time interval T of neuron i is given by

$$C_i = - \sum_{j=0}^N p_j \log_2 p_j \quad (3.1)$$

Using arguments similar to those in Section 2, the representational capacity of n such neurons is $C \approx nC_i$. Note that C has an upper bound of $n \log_2(N+1)$ bits.

We must now express C_i as a function of the average firing frequency of the n neurons. Let the average spike count in the interval T be denoted by μ , i.e., $\mu = E[j]$. We need the discrete probability distribution p_j . The *principle of maximum entropy* provides the most appropriate distribution given the available information (Jaynes 1957). Because we have the mean, the distribution that maximizes the entropy with no additional information about p_j is the geometric (see the Appendix for a proof). The geometric distribution can be written as a function of its mean, μ , as

$$p_j = \frac{\mu^j}{(1+\mu)^{j+1}} \quad (3.2)$$

Substituting the geometric distribution into equation 3.1 yields

$$C_i = -\mu \log_2 \mu + (1+\mu) \log_2(1+\mu) \quad (3.3)$$

where we have let $N \rightarrow \infty$. Note that the functional forms of C_i for the binary and analog cases (equations 2.1 and 3.3) are similar except for the signs in the second term.

We can relate μ to the mean firing frequency in Hz, \bar{f} , by the expression $\mu = \bar{f}T$, which allows C_i to be expressed in terms of the mean firing frequency and the averaging time period as

$$C_i(\bar{f}, T) = -\bar{f}T \log_2(\bar{f}T) + (1+\bar{f}T) \log_2(1+\bar{f}T) \quad (3.4)$$

3.2 Energy Expenditure. Because it takes energy to repolarize the membrane after an action potential, a neuron uses more energy when firing action potentials than when inactive. For the purposes of our first-order estimate here and to be consistent with the binary case, we will assume that the energy needed by a neuron increases linearly with its frequency of firing. Specific energy consumption values are not needed. In this case consider the energy expended at rest relative to the energy expended after an action potential occurs. That is, assume that the energy consumed by a quiescent neuron is one unit and that energy consumption increases linearly with mean firing rate, \bar{f} .

Under this assumption, the energy expended by neuron i can be expressed as $E_i = 1 + k\bar{f}_i$. Assuming that the mean firing rate of each neuron is equivalent to the mean firing rate of the population of neurons, the energy expended by all neurons is

$$E = \sum_{i=1}^n (1 + k\bar{f}_i) = \sum_{i=1}^n (1 + k\bar{f}) = n(1 + k\bar{f}) \quad (3.5)$$

where, as in equation 3.4, \bar{f} denotes the mean firing frequency of all n cells. Dimensional analysis and a comparison of equations 2.2 and 3.5 reveals that $k = (r - 1)/f_N$. Using the same reasoning as in Section 2.2, experimental values of r should range from 10 to 200. As will become apparent in the next subsection, the value of r affects the optimization calculations in a manner similar to the binary case.

3.3 Maximizing C/E . The expression for C/E in the analog neuron case follows from equations 3.4 and 3.5 of the previous two subsections and the approximation $C \approx nC_i$,

$$C/E = \frac{-\bar{f}T \log_2(\bar{f}T) + (1 + \bar{f}T) \log_2(1 + \bar{f}T)}{1 + (r - 1)\bar{f}/f_N} \quad (3.6)$$

and is independent of the number of neurons as in the binary case. Note that the corresponding equation for the binary case, equation 2.3, is similar in structure.

Curves of C/E as a function of \bar{f} with $T = 25$ msec for various values of r are shown in Figure 3 and have shapes similar to those of the binary case shown in Figure 1. The curves in Figure 3 show that a 10-fold increase in the value of r results in more than a 5-fold decrease in the optimal mean frequency. If we set the maximum firing frequency to 400 Hz and let $T = 25$ msec, then $N = 10$. For $r = 10$, the value of \bar{f} that maximizes C/E is 43 Hz; for $r = 200$, the optimal frequency is 6 Hz. These values span a narrower range than the 8–64 Hz range cited in the binary case. Even so, the overlap between the ranges is remarkable, implying that this theory is relatively independent of the type of neural code.

From a physiological viewpoint, we can associate the averaging time with the time constant of the neurons. Because the time constant of neurons is sure to vary and is still a matter of conjecture for animals in their natural environments, we calculated the optimal mean frequencies, as a function of r , for averaging times of 10, 25, and 50 msec. It is apparent from Figure 4 that the averaging time has little effect on the result. Shorter averaging times simply raise the optimal mean frequencies.

4 Summary

Given a population of n binary neurons with np neurons firing on average, the coding scheme with $p = 0.5$ has the largest representational

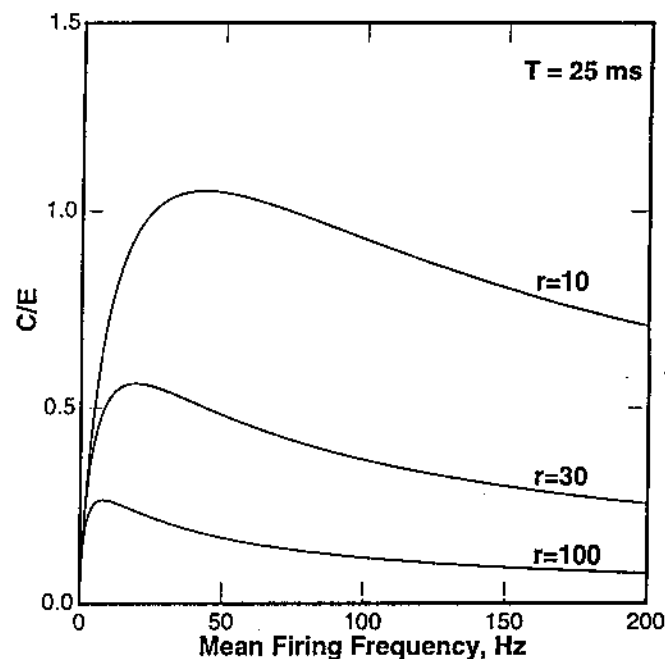


Figure 3: Energy efficiency curves for the analog case. The ratio of representational capacity to energy expenditure (C/E) is a concave function of the average firing frequency (\bar{f}). Here we plot three such curves corresponding to three different values of r , the relative energy expended for an action potential compared to remaining at rest (with an assumed averaging time period of 25 msec and an assumed maximum firing frequency of 400 Hz). Note how the maxima across curves decrease as action potentials become more energetically expensive (i.e., as r increases). These curves are similar in shape to those of the binary case (see Fig. 1).

capacity; that is, on average, half of the neurons are active for any given stimulus. We have shown, however, that this maximum capacity coding scheme is not the most energy efficient scheme. Binary codes that maximize representational capacity and minimize energy expenditure use 2–16% of the number of cells available to fire for any given stimulus, which is considerably less than half of the available neurons.

In the case of analog neurons, the mean firing frequency that maximizes the ratio of the representational capacity to the energy expended is a function of the time interval over which the frequency is determined

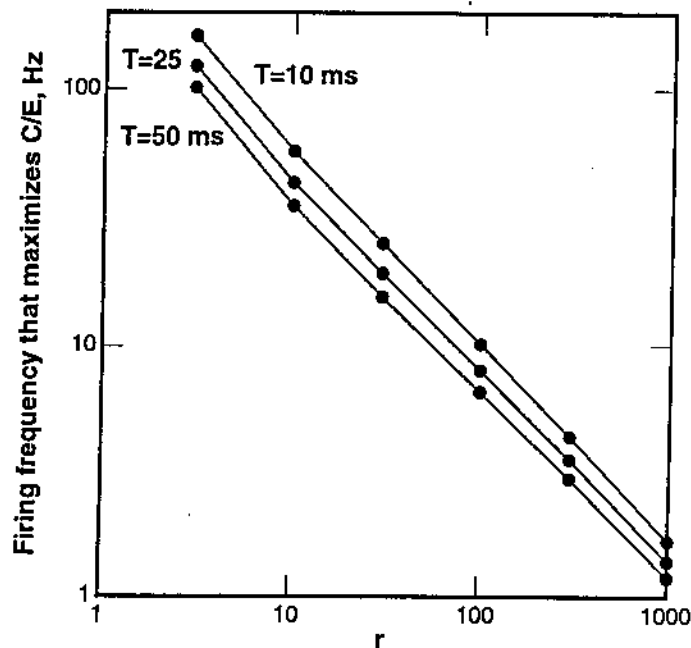


Figure 4: As the relative energy expended for an action potential compared to remaining at rest (r) increases, the optimal firing frequencies for maximizing the ratio of representational capacity to energy expended (C/E) decrease. This decrease is approximately linear on logarithmic scales. The three curves correspond to three different values of the averaging time interval, T . As T increases, the optimal firing frequencies decrease. The filled circles in this graph represent maxima of C/E for different values of r .

and the maximum possible firing frequency. With an averaging time interval of 25 msec and a maximum firing frequency of 400 Hz the optimal mean frequency ranges from 6 to 43 Hz, depending on the metabolic cost of an action potential.

Our results tend to support the notion that as neuronal firing becomes energetically more expensive, lower firing rates become more beneficial. However, our results also show that very low firing rates can be detrimental to energy efficiency. Most importantly, our results support the notion of an energy efficient firing rate. Neurons firing above or below this rate transmit information less efficiently.

Having made these observations relating efficiency (C/E) to the rel-

ative metabolic cost of generating an action potential (r), and having computed firing rates that are efficient, we might compare our results to values actually found in the brain. To make this comparison, it is important to consider a natural situation. Cell firing in anesthetized animals and cell firing triggered by experimenter controlled stimulus presentation seem inappropriate. We are interested in cell firing as an animal behaves more or less naturally in a setting similar to its native environment. Although it is still not possible to measure the activities of a statistically large number of cells simultaneously, open field spatial mapping studies of limbic system activity seem most relevant.

In such studies animals are allowed to wander through the environment and cell firing occurs at each place in the open field. In the rat subiculum (a region of the limbic system) there are two types of frequency firings noted for four types of neurons (Sharp and Green 1994). One type of neuron has an average frequency firing of 9 Hz and is the primary coding element presumably for head direction and where the animal is going. Another class of neurons, presumably interneurons, is characterized by relatively high firing rates (about 37 Hz). Of all cells measured, significantly fewer cells belonged to this second class. Assuming both cell types have similar metabolic costs, this may be interpreted as a confirmation that nature favors energy efficient cells. The observed firing rates of the primary cells fall at the lower end of our computed optimum range while those of the interneurons are near the higher end of our computed values.

It would be remarkable if we were fortunate enough to pick exactly the two parameters that nature has chosen to optimize. Undoubtedly, there are other concerns that nature had in mind when it evolved the frequency firings of various neuronal types. Here, we provide a plausible explanation of why and how the evolution of the brain may have been constrained by energy efficiency as well as statistical dependencies and representational entropies. We have also shown how these constraints relate to observed neural firing frequencies.

Besides enhancing our understanding of the tradeoffs between representational capacity and energy conservation that nature has resolved, this work may be of practical importance. It offers a rationale for designing networks that, instead of maximizing representational capacity, provide an optimal compromise between capacity and energy use. For example, networks with distributed codes that use a large percentage of the available neurons have large representational capacities but may require significant energy, while networks with grandmother codes (e.g., nearest-neighbor networks) generally use less energy but have lower representational capacities. Furthermore, although we have focused on the tradeoff between representational capacity and energy expenditure in biological systems, this optimization approach may also be useful in the design of massively parallel systems where energy consumption and heat dissipation are critical concerns.

Appendix

In this appendix, we provide a proof that the geometric distribution given in equation 3.2 is the distribution that maximizes the entropy. The maximum entropy distribution has the form (Jaynes 1957)

$$p_j = e^{-\lambda_0 - \lambda_1 j} \quad (\text{A.1})$$

with the restriction $\sum_{j=0}^N p_j = 1$. We will assume that $N \rightarrow \infty$.

Let $t = e^{-\lambda_1}$. Then

$$\sum_{j=0}^{\infty} t^j = \frac{1}{1-t} = e^{\lambda_0} \quad (\text{A.2})$$

Since we are given the mean of the distribution, $E[j] = \mu$, we know that

$$\mu = \sum_{j=0}^{\infty} j p_j = (1-t) \sum_{j=0}^{\infty} j t^j \quad (\text{A.3})$$

By taking the derivative of $1/(1-t)$, we obtain the relation

$$\frac{t}{(1-t)^2} = \sum_{j=0}^{\infty} j t^j \quad (\text{A.4})$$

which, upon substitution into the expression for μ , yields

$$\mu = \frac{t}{1-t} \quad (\text{A.5})$$

Solving for t in terms of μ and substituting into the expression for p_j yields the geometric distribution.

Acknowledgments

This research was supported in part by NIH MH00622, MH48161, and Electric Power Research Institute (EPRI) RP8030-08 to W.B.L., Baxter Research, and the Department of Neurological Surgery at the University of Virginia, Dr. John A. Jane, Chairman.

References

- Adelsberger-Mangan, D. M., and Levy, W. B. 1992. Information maintenance and statistical dependence reduction in simple neural networks. *Biol. Cybern.* 67, 469-477.
- Atick, J. J. 1992. Could information theory provide an ecological theory of sensory processing? *Network* 3, 213-251.
- Barlow, H. B. 1959. *Symposium on the Mechanization of Thought Processes*. H. M. Stationary, London, No. 10, pp. 535-539.

- Barlow, H. B. 1969. Trigger features, adaptation and economy of impulses. In *Information Processing in the Nervous System*, K. N. Leibovic, ed., pp. 209-226. Springer-Verlag, New York.
- Cover, T. M., and Thomas, J. A. 1991. *Elements of Information Theory*. John Wiley, New York.
- Foldiak, P. 1990. Forming sparse representations by local anti-hebbian learning. *Biol. Cybern.* 64, 165-170.
- Jaynes, E. T. 1957. Information theory and statistical mechanics I, II. *Phys. Rev.* 106, 620-630; 108, 171-190.
- Kadekaro, M., Crane, A. M., and Sokoloff, L. 1985. Differential effects of electrical stimulation of sciatic nerve on metabolic activity in spinal cord and dorsal root ganglion in the rat. *Proc. Natl. Acad. Sci. U.S.A.* 82, 6010-6013.
- Levy, W. B. 1985. An information/computation theory of hippocampal function. *Soc. Neurosci. Abstr.* 11, 493.
- Mathews, J., and Walker, R. L. 1970. *Mathematical Methods of Physics*, 2nd ed. Benjamin/Cummings, Menlo Park, CA.
- Redlich, A. N. 1993. Supervised factorial learning. *Neural Comp.* 5, 750-766.
- Ritchie, J. M., and Straub, R. W. 1980. Oxygen consumption and phosphate efflux in mammalian non-myelinated nerve fibers. *J. Physiol., London* 304, 109-121.
- Sharp, P. E., and Green, C. 1994. Spatial correlates of firing patterns of single cells in the subiculum of the freely moving rat. *J. Neurosci.* 14(4), 2339-2356.
- Softky, W. R., and Kammen, D. M. 1991. Correlations in high dimensional or asymmetric data sets: Hebbian neuronal processing. *Neural Networks* 4, 337-347.
- Sokoloff, L. 1989. Circulation and energy metabolism of the brain. In *Basic Neurochemistry: Molecular, Cellular, and Medical Aspects*, 4th ed., G. J. Siegel, B. W. Agranoff, R. W. Albers, and P. B. Molinoff, eds., pp. 565-590. Raven Press, New York.
- Yarowsky, P., Kadekaro, M., and Sokoloff, L. 1983. Frequency-dependent activation of glucose utilization in the superior cervical ganglion by electrical stimulation of cervical sympathetic trunk. *Proc. Natl. Acad. Sci. U.S.A.* 80, 4179-4183.
- Yarowsky, P., Crane, A., and Sokoloff, L. 1985. Metabolic activation of specific postsynaptic elements in superior cervical ganglion by antidromic stimulation of external carotid nerve. *Brain Res.* 334, 330-334.

Received January 16, 1995; accepted May 30, 1995.