

Using Energy Efficiency to Make Sense Out of Neural Information Processing

William B Levy¹

University of Virginia Health System
Department of Neurosurgery
Charlottesville, VA
e-mail: wbl@virginia.edu

Robert A. Baxter

Baxter Research
Bedford, MA
e-mail: rob@baxterresearch.com

I. INTRODUCTION

One of the earliest information theoretic interpretations of neural processing and of cognitive processing (e.g., by Barlow and by Watanabe) is recoding to remove statistical dependency while avoiding or minimizing information loss by such a transformation. Since that time there have been many articles that use entropies and relative entropies (including mutual information) to describe various aspects of information in the brain. However, there are two widespread and well known neurobiological observations that cannot be explained by information measures alone.

First, on average binary neuron signaling occurs an order of magnitude below the information optimizing rate. That is, neocortical neurons have a maximal firing rate of 400 Hz, but rather than being used at an average frequency of 200 Hz, which would maximize their information capacity for binary signaling, they have an average frequency of spikes (think of a spike as 1 and a nonspike as 0) of 10-20 Hz. Second, there is a natural form of randomization that is by far the greatest source of noise in neocortical neurons called "quantal synaptic failures." Because of the quantal failure process, excitatory synapses in neocortex are essentially the binary Z-channel of information theory – and what a surprisingly bad Z-channel they are for an optimized system. Under physiological conditions, the failure of an action potential to produce transmitter release (i.e., $1 \rightarrow 0$) is greater than 50%, and often said to be as high as 75%. Consideration of information measures in the context of energy efficiency explains both of these observations.

From the viewpoint of energy consumption, neural processing is rather expensive. The adult human brain accounts for 20% or more of our total energy use, and in young children energy use by the brain can account for nearly 50% of the caloric intake. Thus, there is compelling motivation to hypothesize that microscopic parameterizations of the nervous system are optimized for energy use.

II. ENERGY EFFICIENT AXONAL COMMUNICATION

Reliable transmission between neurons in the brain is mediated by axons. The axon can be considered a binary signaling channel that transmits action potentials (called spikes) from a neuron's cell body to its synapses. In neocortex, axonal transmission is characterized by high signal-to-noise ratios and is nearly perfect. That is, if we call the input to an axon Z and call the output of the axon (after an appropriate time delay) Z' , then we can write $I(Z; Z') = H(Z) - H(Z|Z') \approx H(Z)$ and $H(Z|Z') \approx 0$.

¹Supported by NIH MH48161 and the Meade-Munster Foundation. The authors thank T. Berger for his comments.

Axons consume energy whether or not a spike is conducted. A spike costs about 70 times more than a nonspike over one minimal signaling interval ($1/400$ Hz = 2.5 ms).

The ratio of information transmitted to energy consumed reveals that the neuronal activity (i.e., the average number of spikes emitted per second) can be optimized [1]. That is, calling $E[Z] = P(Z = 1) = p$ and using a constant K to convert to the relative energy use per minimal interval, produces the ratio $H(Z)/[K(1-p)+70p]$, the units of which are bits/joules. Such an optimization produces an optimal average firing rate, p^* , which in this specific case is about 18 Hz. That is, an entropy measure specifically in the context of energy efficient communication explains the observed firing rates of neocortical neurons while entropy considerations alone are well off the mark.

III. SYNAPTIC FAILURES ARE ENERGY EFFICIENT

If we measure neuronal computation and communication as transmission through a capacity-limited channel, then failure rates exist that are inconsequential in terms of information loss and that save energy. Indeed, synaptic failures are very lossy but can be tolerated due to Shannon's source-channel matching results.

Because a neuron receives 10^4 inputs, it could be getting nearly $H(p^*) \cdot 10^4$ bits per 2.5 ms. Yet its output is just a single such axon capable of carrying only $H(p^*)$ bits per 2.5 ms. Thus, large information losses due to synaptic transmission and due to neuronal integration are easily tolerated. Moreover, because the processes of synaptic transmission and postsynaptic excitation are energy consuming, synaptic failures save energy. These observations lead to the following neurocomputational principle: *Maximize the computational information developed by a neuron and its inputs to no more than the limit imposed by the information capacity of the axon whose capacity is set by optimizing the energy efficiency of its signaling.*

We have made use of this principle to prove that there is an optimal synaptic failure rate, $f \approx 4^{-H(p^*)}$. This result is notable for its lack of dependence on number of inputs. It was also somewhat surprising to us that lower values of p^* , or equivalently lower values of $H(p^*)$, imply greater failure rates. More importantly, the values of p^* calculated previously and corresponding to measured values of about 20 Hz ($p = 0.05$ per 2.5 ms) produce failure rates by this formula in the appropriate, empirically observed range, $\sim 70\%$.

REFERENCES

- [1] W. B Levy and R. A. Baxter. "Energy efficient neural codes." *Neural Comp.*, vol. 8, 1996, pp. 531-543.

Using Energy Efficiency to Make Sense Out of Neural Information Processing

William B Levy¹ Member, IEEE and Robert A. Baxter² Member, IEEE

^{1,2}University of Virginia Health System

Department of Neurosurgery

Charlottesville, VA

²Baxter Research Company

Bedford, MA

In attempting to understand neurobiological computation, it is natural to apply – or try to apply – Shannon based information theory. After all, as most every neuroscience textbook says the purpose of the brain is information processing, and there is even an information processing theorem in Shannon's theory. However, going beyond such superficial similarities requires both careful mapping of biology (i.e., neurophysiology and neuroanatomy) into meaningful as well as mathematically tractable units of function and, just as important, some kind of confirming or disconfirming experimental evidence that any particular hypothesis is likely to be true or false.

One of the earliest information theoretic interpretations of neural processing (Barlow, 1959) and of an aspect of cognitive processing (Watanabe, 1969) is recoding to remove statistical dependency while avoiding or minimizing information loss by such a transformation. Since that time there have been many articles that use entropies and relative entropies (including mutual information) to describe various aspects of information in the brain (e.g., Rieke et al., 1999).

Here we point out the importance of considering information processing and information transmission in the context of energy efficiency. In particular, certain qualitative observations that are not explained (i.e., that are not sensible!) when quantified and considered by entropies alone, are made sensible by also considering the energy efficiency of neural communication and neural processing.

The Role of Optimization in Understanding Organisms

The central theory of biology is evolution through natural selection. Naively, that is if we assume a stationary environment, natural selection improves the ability of an average organism to reproduce fertile offsprings. Thus comes the notion of evolution as an optimizing process.

Of course, there are no guarantees that such changes are long-term improvements because the real world is nonstationary. Moreover, recent evolutionary developments – e.g., language in humans – have not been around very long so even if the environment has been stable enough, we would be naïve to assume the human brain is optimized for language (or language for the brain). On the other hand, certain microscopic aspects of the nervous system have been around for much, much longer. That is, the commonality of microscopic features across a set of species that have separately evolved for eons implies that a very good – and we shall hypothesize approximately optimal – solution has been found for neural processing at the microscopic level (i.e., pieces of neurons – axons, dendrites – and small organizations of neurons).

The mammalian line split off from the reptilian line over 250 million years ago, and the various mammalian species that are present today began radiating from their last common ancestor about 200 million years ago. Yet most neurons in mammalian neocortex have the same general properties and not too dissimilar physiological parameterizations, and thus appropriate microscopic optimization under physiological conditions is a sensible hypothesis. Here we point out that there are two such parameters directly related to information transmission and information processing that are, by this hypothesis,

optimized but that simple, textbook information theoretic considerations alone fail to explain. The two facts are:

On average binary neuron signaling occurs an order of magnitude below their information optimizing rate. That is, neocortical neurons have a maximal firing rate of 400 Hz, but rather than being used at an average frequency of 200 Hz, which would maximize their information capacity for binary signaling, they have an average frequency of spikes (think of a spike as 1 and a nonspike as 0) of 10-20 Hz.

Second, there is a natural form of randomization that is by far the greatest source of noise (i.e., noise in the sense of unpredictable random fluctuations) in neocortical neurons called "quantal synaptic failures." Because of the quantal failure process, excitatory synapses in neocortex (but not in some other parts of the brain) are essentially the binary Z-channel of information theory – and what a surprisingly bad Z-channel they are for an optimized system. The probability of a spontaneous quantal synaptic event (i.e., $0 \rightarrow 1$) is nearly zero (experimental data indicates about 10^{-6} per minimal signaling interval). On the other hand, under physiological conditions, the failure of an action potential to produce transmitter release (i.e., $1 \rightarrow 0$) is greater than 50%, and often said to be as high as 75%.

As we shall see, consideration of information measures in the context of energy efficiency explains both of these observations.

Energy Use by the Brain

Organisms in the natural environment are highly efficient in their use of their available energy, and sometimes even more important, their cooling resources. That is, the use of food and water by organisms is often associated with processes that become sensible when we suppose one or both of these resources are scarce. And scarce they are at the margins, the place where evolution occurs fastest. It is at the interfacial niches of marginal survival that competition is most severe and mutations would be most beneficial, thus leading to higher rates of evolution. Clearly at such places of marginal survivability, energy efficiency will be very important. So we have good reason to expect that organisms are made up of and use energy efficient processes.

In this context of energy, neural processing is rather expensive. The adult human brain accounts for 20% or more of our total energy use. It consumes on the order of 20-25 watts. In young children, whose brains are nearly as large as an adults, the energy use by this organ can account for nearly 50% of the caloric intake.

A recent energy audit by Attwell and Laughlin (in press) implies that more than 85% of the energy used by brain goes toward restoring the ion fluxes across neuronal membranes that are the biophysical basis of computation and communication in the neocortex. Thus neural information processing, although perhaps five to six orders of magnitude more energy efficient than man-made computation, is a considerable expense for the organism.

Thus we have further motivation to hypothesize that the microscopic parameterizations of the nervous system are optimized for energy use.

Partitioning Neurophysiological Function

Figure 1 illustrates the partitioning of information transmission, $I(Z, Z')$ and information processing $I(X; \sum \phi(X_i) Q_i)$ which we shall address below. Importantly, and not part of our intuition before we started this work, it is helpful to consider axonal transmission as a separate functional unit, which is distinct from synaptic transmission. Synaptic transmission, itself, is partitioned with neuronal integration.

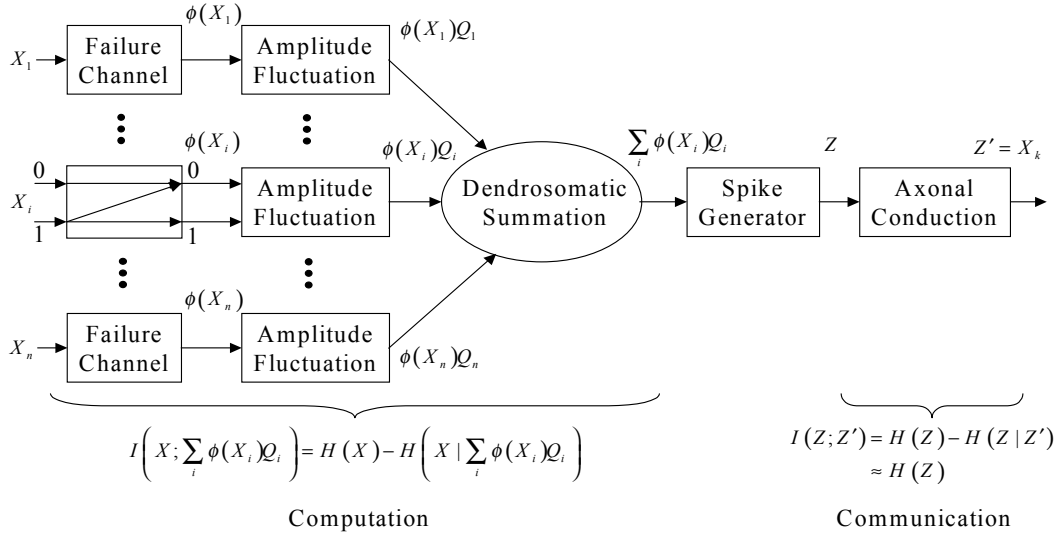


Figure 1. Partitioning communication and computation for a single neuron and its inputs. The presynaptic axonal inputs to the postsynaptic neuron is a multivariate binary vector, $X = [X_1, X_2, \dots, X_n]$. Each input, X_i , is subject to quantal failures (i.e., the Z-channel of information theory), the result of which is denoted by $\phi(X_i)$, another binary vector that is scaled by quantal amplitude fluctuation, Q_i . Thus each input provides excitation $\phi(X_i)Q_i$. The dendrosomatic summation, $\sum_i \phi(X_i)Q_i$ is the endpoint of the computational process, and this sum is the input to the spike generator. The spike generator output is a binary variable, Z , which is faithfully transmitted down the axon as Z' . This Z' is just another X_i elsewhere in the network. In neocortex, experimental evidence indicates that axonal conduction is, essentially, information lossless, as a result $I(Z; Z') \approx H(Z)$. The information transmitted through synapses and dendrosomatic summation is the mutual information $I(X; \sum_i \phi(X_i)Q_i) = H(X) - H(X \mid \sum_i \phi(X_i)Q_i)$.

Here we only consider a simple form of neuronal integration – adding up of exclusively excitatory inputs; the numerically inferior inhibitory inputs are ignored. We avoid the problem of the spike generator for now. Finally it should be obvious that any conclusions drawn about Z' of Fig. 1 apply to the X_i 's since they are generically identical.

Optimization of Axonal Communication in the Context of Energy Efficiency (Levy & Baxter, 1996)

Long-distance transmission in the brain (i.e., distances exceeding ~ 0.5 mm and often shorter) is mediated by a part of a neuron called the axon. The axon is a binary signaling device that transmits spike-like pulses (called action potentials) from a neuron's cell body to its synapses. In neocortex the axon runs at such high signal to noise levels that we may as well assume it is perfect. That is, if we call the input to an axon, originated by the decision-making portion of a neuron called the initial segment, Z , and call the output of the axon at the appropriate time delay Z' , then we can write

$$I(Z; Z') = H(Z) - H(Z \mid Z') \approx H(Z). \text{ That is,}$$

$$H(Z \mid Z') \approx 0.$$

Such an axon is capable of emitting a spike or nonspike as fast as 400 times per second. With statistical independence between such binary messages, the channel capacity would be 400 bits/sec if the average

spike rate were 200 Hz. However, when such an optimization argument meets neurophysiology data, this prediction is far, far off. With the exception of certain neurons closer to motor control, the typical neocortical neuron has an average firing frequency of 10-20 Hz. Clearly information transmission is not being optimized. However, let us bring energy considerations into the mix.

Axons consume energy even when not conducting a spike. Just transmitting zero's (i.e., nonspikes) costs. This cost (actually the relative cost) has been measured in various odd places. The best data indicate that a spike costs about 70 times more than a nonspike over one minimal signaling interval (1/400 Hz = 2.5 ms). Using such measurements and hypothesizing a ratio that is essentially bits/joule, leads to a satisfactory result in terms of an optimization. That is, calling $E[Z] = P(Z = 1) = p$ and using a constant K to convert the relative energy use per minimal interval to joules, produces the ratio

$$\frac{H(Z)}{K((1-p) + 70p)}$$

which is bits/2.5 ms divided by joules per 2.5 ms.

This ratio is concave (cap), and when maximized across p , the constant K disappears. Such an optimization produces an average firing rate, p^* , which in this specific case is about 18 Hz. That is, an entropy measure specifically in the context of energy efficient communication explains the observed firing rates of neocortical neurons while entropy considerations alone are well off the mark.

Synaptic Failures are Energy Efficient

As pointed out above, synaptic transmission is extremely noisy. However, if we measure neuronal computation as if it were a channel, then there exists failure rates (Z crossover probability $1 \rightarrow 0$) that are inconsequential in terms of information loss and that save energy. Indeed, both neuronal computation (most simply, a neuron adding up its inputs) and synaptic failures are very lossy but can be tolerated due to Shannon's source-channel matching results (Shannon, 1948).

Because a neuron receives $\sim 10^4$ inputs, it could be getting nearly $H(p^*) \cdot 10^4$ bits per 2.5 ms. Yet its output is just a single such axon capable of carrying only $H(p^*)$ bits per 2.5 ms. Thus, large information losses due to synaptic transmission and due to neuronal integration are easily tolerated. Moreover, because the processes of synaptic transmission and postsynaptic excitation are energy consuming, synaptic failures save energy. These observations lead to the following conjectured principle governing neuronal computation.

Maximize the computational information developed by a neuron and its inputs to no more than the limit imposed by the information capacity of the axon whose capacity is set by optimizing the energy efficiency of its signaling.

Here are the results of a formal analysis to support these statements.

Assumptions:

A0: A computation by an excitatory neuron is the summation of its inputs every computational interval. The mutual information of such information processing is closely approximated as

$$I^C = I(X; \sum \phi(X_i)Q_i)$$

where X is an n -dimensional binary input to a neuron.

A1: Axons are binary signaling devices carrying independent spikes and used at their energy optimum; that is, each axon is used at the information rate C^E bits per computational interval which implies firing probability p^* .

A2: The number of inputs to a neuron is not too small – say $n > 2/p^*$. Clearly this is true in neocortex.

A3: With the proviso that A1 and A2 must be obeyed, a process requiring less energy is preferred to a process requiring more energy.

A4: The spike generator at the initial segment creates a bitwise code suitable for the axonal channel, and this encoding is nearly perfect in using the information received from the dendrosomatic computation. That is, as an information source the spike generator produces information at a rate of nearly $H(p^*)$.

From these assumptions we have a lemma.

Lemma 1: $I^C \geq H(p^*)$. That is, the only way to use an axon at its energy optimal rate, $H(p^*)$, is to provide at least that much information to it for possible transmission. In particular, this lemma dictates: quantal failures, as an energy saving device, will be used (or failure rates will only be increased) only when I^C is strictly greater than $H(p^*)$.

This lemma and a specific biophysical model containing an appropriate region of monotonic dendrosomatic summation per joule cost leads to two theorems concerning an optimal failure rate. The averaged summed postsynaptic activation, $E\left[\sum_{i=1}^n X_i\right]$, or more accurately $E\left[\sum \phi(X_i)Q_i\right]$, should be as small as possible because of (A3) while (A2) maintains n and while (A1) maintains p^* . This restricted minimization of average synaptic activation implies processes, including synaptic failures, that reduce energy use. But when operating on the energy efficient side of the depolarization versus information curve, reducing the average summed activation monotonically reduces I^C , and this reduction is unrestricted until lemma 1 takes force. That is, this reduction of I^C should go as far as possible because of (A3) but no lower than $H(p^*)$ because of the lemma. As a result, energy optimal computation is characterized by

$$I^C = H(p^*) = C^E,$$

an equality which we call **Theorem G**. Accepting Theorem G leads to the following corollary about synaptic failures:

Corollary F: Provided $np^* > 2$, neuronal computation is made more energy efficient by a process of random synaptic failures. In particular, there is an optimal failure f ,

$$f \approx \left(\frac{1}{4}\right)^{H(p^*)}.$$

Proof that $I\left(X; \sum_i \phi(X_i)\right) = H(p^*) \Rightarrow f \approx \left(\frac{1}{4}\right)^{H(p^*)}$: Call $s = 1 - f$ and n the number of inputs.

We will use $I\left(X; \sum_i \phi(X_i)\right) = H\left(\sum \phi(X_i)\right) - H\left(\sum \phi(X_i) | X\right)$.

Because the synaptic failures are conditionally independent, $H\left(\sum \phi(X_i) | X\right) = H\left(\sum \phi(X_i) | \sum X_i\right)$,

and $P\left(\sum \phi(X_i) = y^\phi | \sum X_i = y\right) = \frac{y!}{(y - y^\phi)! y^{\phi!}} s^{y^\phi} (1 - s)^{y - y^\phi}$.

Because n is so large, normal distributions are well-approximated for both of the required entropies; i.e.,

$$I(\sum X_i; \sum \phi(X_i)) \approx \frac{1}{2} \log 2\pi n p s - E_y \left[\frac{1}{2} \log 2\pi e y s (1-s) \right].$$

But this large n and the normal approximation also imply: $E[\log y] \approx \log E[y] = \log np$. Therefore

$$I(\sum X_i | \sum \phi(X_i)) \approx -\frac{1}{2} \log_2 (1-s) = -\frac{1}{2} \log_2 f, \text{ and as Theorem G requires } H(p^*) \text{ to equal this}$$

mutual information, $f \approx \left(\frac{1}{4}\right)^{H(p^*)}$.

Comment: Q_i is another independent source of random fluctuation that is part of synaptic transmission. A similar, but more complicated, derivation can include this variable. Using observed values of the mean and variance of Q_i , this calculation implies that ignoring Q_i produces less than a 10% change in our estimate of f .

This result, $f \approx \left(\frac{1}{4}\right)^{H(p^*)}$, is notable for its lack of dependence on number of inputs. It was also

somewhat surprising to us that lower values of p^* , or equivalently lower values of $H(p^*)$, imply greater failure rates. More importantly, the values of p^* calculated previously and corresponding to measured values of about 20 Hz ($p = 0.05$ per 2.5 ms) produce failure rates by this formula in the appropriate, empirically observed range, $\sim 70\%$.

Why this energy-efficient information processing approach should work so well in predicting (i.e., interrelating) seemingly independent physiological processes is curious because several other constraining optimizations might, *a priori*, seem just as reasonable. For example, constraints on memory capacity, processing speed, and size of the brain have been discussed in the literature. Perhaps we will eventually find that Nature has optimized several constraints simultaneously, and all of these optimizations imply the same microscopic parameterizations.

Acknowledgements

This work was supported by NIH MH48161 to WBL and the Meade-Munster Foundation. The authors thank Toby Berger for his useful comments that helped improve an earlier version of this manuscript.

References

- Atwell, D. & Laughlin, S. B. An energy budget for signalling in the grey matter of the brain. *J. Cerebral Blood Flow & Metabolism*, 2001, in press.
- Barlow, H. B. Symposium on the Mechanization of Thought Processes. London: H. M. Stationary, No. 10, pp. 535-539, 1959.
- Levy, W. B & Baxter, R. A. Energy efficient neural codes. *Neural Comp.* 8, 1996, 531-543.
- Rieke, F., Warland, D., & Bialek, W. *Spikes: Exploring the Neural Code*. Cambridge: MIT Press, 1999.
- Shannon, C. E. (1948) A mathematical theory of communication. *Bell System Tech. J.* 27, 379-423, 623--656).
- Watanabe, S. *Knowing and Guessing; A Quantitative Study of Inference and Information*. New York: Wiley, 1969.