Another Contribution by Synaptic Failures to Energy Efficient Processing by Neurons

Joanna Tyrcha[1*], William B Levy[2]

[1]Dept. of Mathematics, Univ. of Stockholm, Stockholm, Sweden; [2]Department of Neurosurgery, Univ. of Virginia, Charlottesville, VA; [*]Corresponding author

Keywords: skewness, kurtosis, neural network, central limit theorem, quantal transmission

Abstract. Energy efficient use of axons requires achieving a prespecified firing rate. We hypothesize that the failure process of quantal synaptic transmission helps a neuron better approximate this desirable firing rate by moving the neuron's input excitation distribution closer to a gaussian. If there are many statistically independent inputs per neuron, quantal failures do not help but, essentially, are harmless for achieving the desired firing rate. However, such statistical independence is unrealistic. An input distribution reflecting statistical dependence is a mixture distribution. For certain mixtures, failures can improve the gaussian approximation and more precisely produce the desired firing rate.

1. INTRODUCTION

The research described here is part of a general search for positive roles played by the random failure of quantal synaptic transmission. That is, random synaptic failures lose information so they are ostensibly harmful to information processing. However, such synaptic information loss may have little effect on the information eventually transmitted by the postsynaptic neuron while, at the same time, failures will save energy (Levy and Baxter, 2002). In order to quantify this last idea, we took advantage of assumptions that imply the summed synaptic excitation of the postsynaptic neuron obeys a central limit theorem (c.l.t.), and as a result, the entropy of the excitation was well approximated by the entropy of a gaussian distribution. In addition to analytical convenience, the c.l.t. may be used implicitly by a neuron to achieve its energy-efficient firing rate (see also Levy, 2003). Here the mechanism proposed for setting firing rate is simple: a neuron has kept track of the mean and the variance of the excitation induced by its inputs, and threshold is set, implicitly, by some biophysical mechanism that is equivalent to the neuron assuming a gaussian distribution.

In most brain regions, a hallmark of neuronal computing is the summation of a high dimensional input at each postsynaptic neuron. Because the number of inputs is finite, because the contribution of each input is finite and because there are so many inputs being summed, we have every reason but one to invoke a central limit theorem. In particular, the statistical dependence of the input could prevent the application of a c.l.t. Here we define some conditions in which quantal failure improve the gaussian approximation.

2. FAILURE PROCESS

In the hippocampus and in neocortex, excitatory synaptic connections dominate and are remarkably unreliable. Each synapse transmits, at most, a single standardized package called a quantum ($\sim 10^4$ neurotransmitter molecules). When an action potential arrives presynaptically, the probability of evoking the release of one such quantal package is reported to range from 0.25 to 0.9 with 0.25 being quite common (Thomson, 2000), especially when one takes into account the spontaneous rates of neurons (Stevens and Wang, 1994; Destexhe and Paré, 1999). The failure of quantal synaptic transmission is a random process (Katz, 1966), and because it loses information, it is counterintuitive when it exists under physiological conditions. Although the failure process causes information loss, combining energetic and informational constraints leads to an optimal, nonzero failure rate for an information-transforming neuron (Levy and Baxter, 2002). In this paper we investigate the quantal failure process for a simple information transforming neuron like a stellate cell of layer IV in neocortex. Using a McCulloch and Pitts model, the neuron adds up its inputs and if this sum is above a threshold value, it fires (produces a one). If below this threshold, nothing (a zero) is transmitted down its axon. To this McCulloch-Pitts mechanism of summation and fire, we add the quantal failure process. The quantal failure process is seen to be equivalent to the so-called Z-channel of information theory well-studied (e.g., Gallagher (1968)) in terms of the information it loses. Such synaptic transmission is a stochastic process that, for input *i*, leaves

$X_i = 0$ unchanged, but it can change $X_i = 1$ to 0 via a failed transmission, where such failures occur at rate $f \in [0,1)$ and occur independently across all *i*. Equivalently, successful transmission is a Bernoulli process governed by $s = 1 - f$, the probability of success. The summed excitation the neuron receives is a random variable $Y = \sum X_i$, and as the crudest approximation suppose its inputs are binomially distributed (*Bin*(*n*, *p*)). If the quantal failure process is operating, the summation over this set of excitatory inputs, $Y^* = \sum X_i^*$, is also binomial distributed but with parameters *n* and *sp*. (The asterisk notation will denote that the quantal failure process is operating.)

3. GAUSSIAN APPROXIMATION

Lets first assume a simple input, i.e., full independence between equiprobable inputs that implies a binomial distribution. Because there are many inputs, we approximate the binomial distribution by a gaussian. For a binomial distribution with parameters *n* and *p*, the expected value of excitation equals *np* and the variance is equal to *np(1-p)*. Approximating a binomial distribution by a gaussian one, we can calculate a probability that the summation of inputs is greater than a particular threshold as follows

$$P\left(\sum X_i > \theta\right) = P\left(\frac{\sum X_i - np}{\sqrt{np(1-p)}} > \frac{\theta - np}{\sqrt{np(1-p}}\right) = 1 - \Phi\left(\frac{\theta - np}{\sqrt{np(1-p)}}\right),$$

where Φ is a cumulative distribution function in the standardized normal distribution.

This value of probability is the average firing rate of the neuron per computational interval. If the desired value is equal to 5%, the neuron can set its threshold as $\theta = 1.6449 \cdot \sqrt{np(1-p)} + np$ and, thus, achieve the appropriate firing rate. But if the failure process operates, $\theta^* = 1.6449 \cdot \sqrt{nsp(1-sp)} + nsp$. To see how the failure process affects the gaussian-based threshold we ran simulations. We generated 8000 random numbers for a variety binomial parameterization. For each such simulated distribution, we calculated the 95% quantile for the true distribution (Q_{Best}) and the 95% quantile for the gaussian distribution (Q_{Gauss}), using the mean and variance of the simulation. The accompanying table summarizes four of the different simulations.

Independent Inputs and a Desired Firing Rate of 5%									
		s = 1					*s = 0.2*		
Input prob. p		Q_{Best}	Q_{Gauss}	Firing rate based on Q_{Gauss}		Q_{Best}	Q_{Gauss}	Firing rate based on Q_{Gauss}
0.05	a)	433	432.478	0.049	c)	95	94.537	0.055
0.06	b)	515	515.083	0.049	d)	113	112.057	0.048
(n = 8000 inputs for all examples)								

In these simulations, the gaussian-based threshold (the rounded version of Q_{Gauss}) leads to firing rates very close to the desired rate of 5% and not much different than the optimal threshold, Q_{Best}. Results for 10% firing rates are a little less accurate but still very good. In general quantal failures lead to less accurate firing rates, but this inaccuracy is slight.

4. EFFECT OF QUANTAL FAILURE PROCESS ON SKEWNESS AND KURTOSIS

Continuing with this binomial excitation, we can explain how the quantal failure process affects the skewness and kurtosis of such an input distribution. Because the skewness and kurtosis of a gaussian are known, we can see exactly how s changes these characterizations relative to the gaussian.

The skewness in the binomial distribution $Bin(n, p)$ is equal to $\frac{1-2p}{\sqrt{np(1-p)}}$ and, if the failure process operates, it is $\frac{1-2sp}{\sqrt{nsp(1-sp)}}$. Because the skewness of a gaussian distribution equals zero, a binomial distribution is better fit by a gaussian when the skewness is closer to zero. For the best approximation, it then follows that $1 - 2sp = 0 \Rightarrow sp = 1/2$. Thus, skewness grows worse with quantal failures when $0 < p < 1/2$, and this is the biologically relevant range of p. So quantal failures do not fit our hypothesis when there is binomial excitation and when the third moment is considered. For kurtosis we come to a similar conclusion although the effect seems smaller.

The kurtosis of a binomial distribution is equal to $3 - 6/n + 1/(np(1-p))$ and for the gaussian distribution kurtosis equals 3. So to evaluate effects of kurtosis, we set these two equal and are lead to consider the roots of the equation $p^2 - p + 1/6 = 0$ implying $p = 0.21$ and $p = 0.79$. When $sp = 0.21$, quantal failures helps approximate a gaussian by getting kurtosis right. However, this is only true when p is larger than 0.21. Thus, when $p < 0.21$ and s < 1, kurtosis moves away from the gaussian value. So we can conclude that the failure process hurts when $p < 0.21$, a value of p which is arguably the usual biological condition (firing rates in neocortex are $p \approx 0.12$ per 2.5 msec).

5. MIXTURE INPUTS

The binomial distributions just examined are not very appealing as a model of the input to a neuron. The statistical independence between inputs, which the binomial distribution implies, is unrealistic and trivializes the problems of prediction and retrodiction for neuron based computation. A more relevant model assumes the inputs are sampled from a mixture distribution. Indeed, systems of mixture distributions called hidden markov models are claimed to have the potential complexity to model anything in the world. Therefore, we now examine the effect of quantal synaptic failures on an input model that is a mixture. For purposes of simplicity, we discuss the mixture of two binomial distributions:

$$p(k \mid n, p_1, p_2, w) \overset{def}{=} P(\Sigma X_i = k) = w\binom{n}{k} p_1^k (1-p_1)^{n-k} + (1-w)\binom{n}{k} p_2^k (1-p_2)^{n-k}$$

where $0 < w < 1$ is the so-called mixing parameter.

For the excitation produced by such a two component mixture, we can calculate, exactly, their skewness and kurtosis. First note that if X is binomially distributed with generic parameters n and p, its first four moments are

$$E[X] = np;\ E[X]^2 = np + n(n-1)p^2;$$
$$E[X]^3 = np + 3n(n-1)p^2 + n(n-1)(n-2)p^3;$$
$$E[X]^4 = np + 7n(n-1)p^2 + 6n(n-1)(n-2)p^3 + n(n-1)(n-2)(n-3)p^4.$$

The same moments for the mixture of two binomial distributions are combinations of moments for distributions in the mixture; for example,

$$w\left[np_1 + n(n-1)p_1^2\right] + (1-w)\left[np_2 + n(n-1)p_2^2\right]$$

is the second moment for the mixture. Likewise, expanding the definitions of skewness and kurtosis, skewness = E[X − E[X]]^3 / σ^3, and kurtosis = E[X − E[X]]^4 / σ^4,

we can calculate the skewness and the kurtosis for the mixture. The resulting equations are quite complicated and will not be shown here.

To simulate the two component mixtures, we let the number of inputs, n, be 8000, and the desired firing rate 0.05. Using the two component mixture formulation, we examine both a positive skewness ($w = 0.7$) and a negative skewness ($w = 0.3$) example with $p_1 = 0.05$ and $p_2 = 0.06$.

The mixtures are illustrated in Fig. 1 with results as follows:

(a) positive skewness, no failure: Q_{Best} = 500, Q_{Gauss} = 492.22; skewness=0.63, kurtosis=2.33

(b) negative skewness, no failure: Q_{Best} = 511, Q_{Gauss} = 525.33; skewness= -0.52, kurtosis=2.23

Note how far apart the gaussian-based thresholds (Q_{Gauss}) are from the best thresholds (Q_{Best}).

(INSERT FIG. 1 ABOUT HERE.)

Now we allow the failure process to operate and consider mixtures with the same parameters as before but multiply each p_i by s = 0.2. Then the results are:

(c) positive skewness, failures=0.8: Q_{Best} = 106, Q_{Gauss} = 103.97; skewness=0.37, kurtosis=2.98

(d) negative skewness, failures=0.8: Q_{Best} = 110, Q_{Gauss} = 110.94; skewness=-0.06, kurtosis=2.82

Note the improvement concerning the difference between both thresholds. In addition to Q_{Gauss} and Q_{Best} moving closer together, even more can be seen in the Fig. 1 plots of these simulated excitations; the failure process can convert an inherently bimodal mixture distribution to an inherently unimodal distribution. Likewise in both cases, quantal failures produce a closer approximation to a gaussian distribution in that skewness moves towards zero and kurtosis moves towards the gaussian value of three.

6. ANALYSES

To understand when quantal synaptic failures will decrease the absolute value of skewness, we calculated the derivative of skewness with respect to s. It is

$$\frac{d(skewness)}{ds} = \frac{d\,\mu_3}{ds} \cdot \frac{1}{var^{3/2}} - \frac{3}{2} \cdot \frac{\mu_3}{var^{5/2}} \cdot \frac{d\,var}{ds}$$

where μ_3 and var are, respectively, the third central moment and variance of the mixture of two binomial distributions with sp_1 and sp_2. There is an important relationship between the change in skewness and failure process. We can restrict our attention to $1-f=s=1$ versus $s=1-\delta$ (δ is small and positive). The derivative of skewness at $s=1$ is

$$\frac{d(skewness)}{ds_{s=1}} = \frac{3}{2} \cdot mean \cdot \frac{1}{var} \cdot \left(-\frac{2\sqrt{var}}{mean} + \frac{2}{3\sqrt{var}} + skewness \right)$$

where $mean$ denotes the expected value of the mixture of two binomial distributions. To use this result, cases of positive and negative skewness must be considered separately.

Quantal failures can reduce skewness in two cases (each with a pair of conditions to satisfy):

Case 1:

$$skewness > 0,\ and\ \frac{d(skewness)}{ds} > 0,\ or\ equivalently,\ skewness > 0,\ and\ \frac{2\sqrt{var}}{mean} - \frac{2}{3\sqrt{var}} > 0.$$

However in this case, skewness cannot be driven to zero.

There is a minimum positive valued skewness for some $0 < s < 1$ (since $\lim_{s \to 0} (skewness) = \infty$), and if skewness is a continuous function of s, then there must be an inflection point between *0* and *1*.

(INSERT FIG. 2 ABOUT HERE.)

Case 2: $skewness < 0\ \ and\ \ \frac{d(skewness)}{ds} < 0,\ \ or\ \ skewness < 0\ \ and\ \ \frac{2\sqrt{var}}{mean} - \frac{2}{3\sqrt{var}} < 0.$

Moreover in this case, skewness can always be driven to zero.

That is, as $s \to 0$, $skewness > 0$ although at $s = 1$ $skewness < 0$. Figure 2 illustrates the distinction between the two cases for a pair of specific examples.

Case 3: *There is no reduction in absolute value of skewness if neither Case 1 nor Case 2 conditions hold for the sign of skewness and its derivative.*

In conclusion, when n is large and the inputs are statistically independent, quantal failures do little harm to gaussian-based thresholding to achieve its energy-efficient firing rate. More importantly where statistical independence does not hold, quantal failures will, in many cases, improve the approximation.

ACKNOWLEDGEMENTS

This work was supported by NIH MH48161, MH63855 and RR15205 to WBL.

REFERENCES

W. B Levy and D. M. Adelsberger-Mangan Is statistical independence a proper goal for neural network preprocessors? INNS World Congress on Neural Networks (1995) I-527-531.

A. Destexhe and D. Paré Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo, J. Neurophysiol. 81 (1999) 1531-1547.

R. G. Gallagher Information Theory and Reliable Communication. (New York, NY, 1968).

B. Katz Nerve, Muscle, and Synapse. (New York, NY, 1966).

W. B Levy Contrasting rules for synaptogenesis, modification of existing synapses, and synaptic removal as a function of neuronal computation, unpublished (CNS03 submission).

W. B Levy and R. A Baxter Energy efficient neural codes, Neural Comp. 8, 1996, 531-543.

W. B Levy and R. A. Baxter Energy-Efficient Neuronal Computation Via Quantal Synaptic Failures, J. Neurosci. 22 (2002) 4746-4755.

C. F. Stevens and Y. Wang Changes in reliability of synaptic function as a mechanism for plasticity, Nature 371 (1994) 704-707.

A. Thomson Facilitation, augmentation and potentiation at central synapses, Trends Neurosci. 23 (2000) 305-312.

FIGURE LEGENDS

Figure 1. Failures can reduce skewness and improve gaussian-based threshold valuation. In these two examples, random synaptic failures move both negative skew (A vs. C) and positive skew (B vs. D) closer to the gaussian value of zero. All four plots show the excitation histograms of a two component mixture, with component one ~ Bin (8000, sp_1) and component two ~ Bin (8000, sp_2) where $p_1 = 0.05$ and $p_2 = 0.06$. In (A) and (C), $w = 0.3$ while in (B) and (D) the mixing probabilities are reversed, i.e., $w = 0.7$. The failure rate in (A) and (B) is zero (i.e., $s = 1$) while in (C) and (D) the failure rate is 80% (i.e., $s = 0.2$). Relative to a desired 5% average firing rate, Q_{Gauss} is the threshold based on a gaussian assumption and is demarcated by a thick vertical line; the actual setting needed to produce as near as possible to 5% is called Q_{Best} and is demarcated by a thin vertical line. Note how these two lines get closer when s = 0.2 compared to $s = 1.0$. Data are based on eight thousand samplings.

Figure 2. Examples of Case 1 and Case 2. Quantal failures will decrease absolute skewness when either Case 1 or Case 2 conditions hold. For Case 1 however, there is no failure rate that will produce zero skew; the minimum, nonzero skewness value occurs at $s = 0.071$. For Case 2, there will always be a failure rate producing zero skewness (in this case when $s = 0.158$). For these simulations: $n = 8000$, $p_1 = 0.05$, $p_2 = 0.06$; in Case 1, $w = 0.7$; in Case 2, $w = 0.3$.

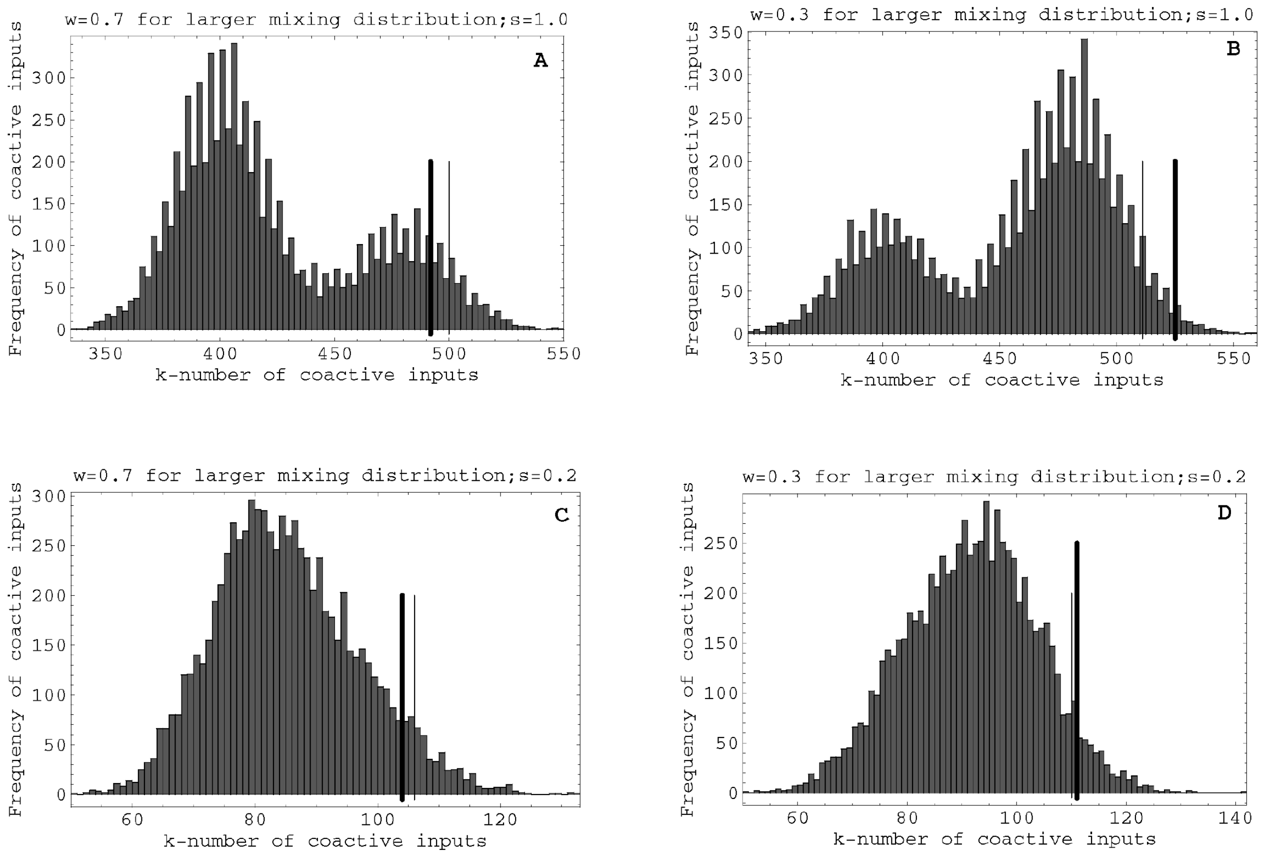

Figure 1

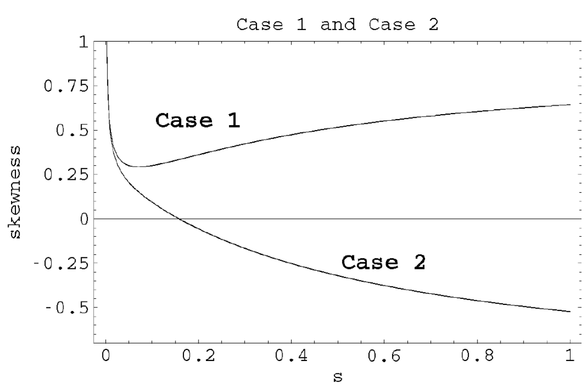

Figure 2