Another Contribution by Synaptic Failures to Energy Efficient Processing by Neurons

Joanna Tyrcha

Department of Mathematics, University of Stockholm, 106 91 Stockholm, Sweden

William B Levy

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

Correspondence should be sent to:

Dr. W. B Levy

University of Virginia Health System

P.O. Box 800420 Neurosurgery

Charlottesville, VA 22908-0420, USA

Phone: 434-924-5014, Fax 434-982-3829, e-mail wbl@virginia.edu

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

Abstract. Energy efficient use of axons requires achieving a prespecified firing rate. We hypothesize that the failure process of quantal synaptic transmission can help a neuron better approximate this desirable firing rate by moving its input excitation distribution closer to a gaussian distribution. We show that quantal synaptic failures essentially are harmless for setting accurate firing rates when there are many statistically independent inputs per neuron. However, a statistically independent set of inputs is unrealistic. A more generic input distribution reflecting input statistical dependence is a mixture distribution. For mixtures, failures can improve the approximation of the desired firing rate via improving the gaussian approximation.

I. 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 implied the summed synaptic excitation of the postsynaptic neuron obeyed 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 neurons to achieve their energy-efficient firing rate (see also Levy, this conference). Here the mechanism proposed for setting firing rate is simple: a neuron keeps track of the mean and the variance of the excitation induced by its inputs. Threshold is set, implicitly, by some biophysical mechanism that is equivalent to the neuron assuming a gaussian distribution. That is, based on the mean and variance of the excitation, a gaussian distribution can be inverted to specify the threshold value (quantile) that gives the energy-efficient firing rate. But if this is true, then we must wonder about the validity of the c.l.t. when applied to a neuron's excitation.

In the neocortex, and most other 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, which implies a gaussian distribution of the summed input excitation.

In particular, the statistical dependence of the input could prevent the application of c.l.t. but this seems very unlikely as there are so many inputs that are far from highly correlated. More likely however, and of some concern, the rate of convergence will be affected by the statistical dependence of the inputs. This paper begins to understand and to quantify the effects of synaptic failures on the excitation of a neuron and the distributional form of this excitation.

II. 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 (~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 information 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 granule cell of the dentate gyrus or a stellate cell of layer IV in neocortex. We model such a neuron as McCulloch and Pitts did. 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 pictured in Fig. 1 and is seen to be equivalent to the so-called Z-channel of information theory (of course, it is well-known (e.g., Gallagher (1968)) that the Z-channel loses information). Such synaptic transmission is a stochastic process that leaves $X_i = 0$ unchanged, but it can change $X_i = 1$ to $X_i = 0$ via a failed transmission, where such failures occur at rate $f \in [0,1)$ and independently across all *i*. Equivalently, successful transmission is a Bernoulli process governed by $s = 1 - f$, the probability of success. The asterisk notation will denote that the quantal failure process is operating (for example $\sum X_i^* = Y^*$).

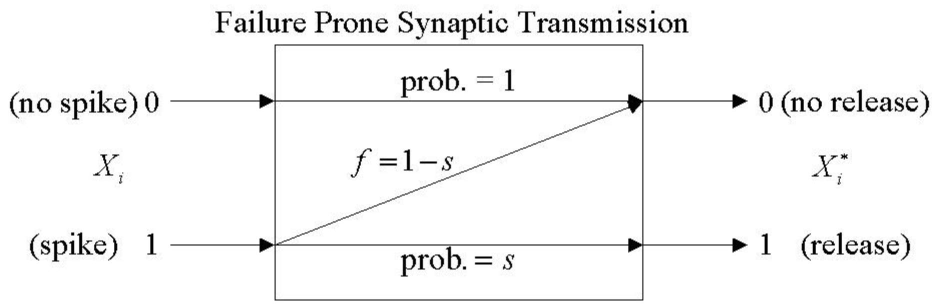

Figure 1. At a single neocortical or hippocampal synapse, excitatory synaptic transmission is assumed to be the Z-channel of information theory. One input state ($X_{i=0}$, in the biological situation at excitatory neocortical synapses) is transmitted perfectly. The other input state ($X_i = 1$) relies on a Bernoulli process to generate the output $X_i^* = 1$.

So, the summation over a particular set of excitatory inputs that the neuron receives is a random variable $Y = \sum X_i$, and as the crudest approximation we might suppose the inputs are binomially distributed (*Bin*(n, p)). If the quantal failure process is operating the summation over a set of excitatory inputs $Y^* = \sum X_i^*$ is also binomial distributed but with parameters n and sp. Let us see why Y^* has a binomial distribution with one of parameters equals sp.

According to our definition of the failure process

$$s \overset{def}{=} P(X_i^* = 1 \mid X_i = 1) = \frac{P(X_i^* = 1 \;\; and \;\; X_i = 1)}{P(X_i = 1)} = \frac{P(X_i^* = 1)}{P(X_i = 1)} = \frac{P(X_i^* = 1)}{p}$$

where the last step follows because X_i^* implies X_i in this case. And then equating s to the last expression, we get $P(X_i^* = 1) = sp$.

Moreover

$$P(X_i^* = 0) = P(X_i^* = 0 \mid X_i = 1) \cdot P(X_i = 1) + P(X_i^* = 0 \mid X_i = 0) \cdot P(X_i = 0) =$$
$$= (1-s)p + 1 \cdot (1-p) = p - sp + 1 - p = 1 - sp.$$

It follows that X_i^* is a Bernoulli random variable with parameter sp, so $Y^* = \sum X_i^*$ is, from the definition, binomial distributed (*Bin*(n, sp)).

Now recall our claim that there is an optimal average firing rate (Levy and Baxter, 1996); then let us ask how can a neuron achieve this optimal firing rate?

We hypothesize that a neuron adjusts its threshold as if its excitation has gaussian distribution. That is, a neuron behaves as if the central limit theorem is operating, and thus, it uses the mean and variance of its postsynaptic excitation to produce a threshold. However, relative to the desired, energy-efficient firing rate, the accuracy of the actual firing rate clearly hinges on the gaussian distribution approximating the distribution of the neuron's internal, summed synaptic excitation.

III. GAUSSIAN APPROXIMATION

Lets first assume a simple input, i.e., full independence between inputs which implies a binomial distribution. Because there are so many inputs being summed, we can 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

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

This value of probability is the average firing rate of the neuron per computational interval. So, assuming that the desired value is equal to 5%, the neuron can set its threshold as

$\theta = 1.6449 \cdot \sqrt{np(1-p)} + np$ and achieve the appropriate firing rate.

As we mentioned above, if the quantal failure process operates, the summation of inputs is also binomial distributed but with parameters *n* and *sp*. Once again, assuming an approximation by a gaussian distribution, the neuron would set the threshold in the similar way, i.e.

$$\theta^* = 1.6449 \cdot \sqrt{nsp(1-sp)} + nsp.$$

To see how the quantal failure process affects the threshold we ran some simulations.

We generated 2000 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. We got the following results:

Independent Inputs and a Desired Firing Rate of 5%								
	s = 1				*s = 0.3*			
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)	116	116.406	0.048	d)	39	38.832	0.063
0.15	b)	326	325.044	0.056	e)	105	104.615	0.061
0.25	c)	532	531.524	0.053	f)	170	168.66	0.060
(n = 2000 inputs for all examples)								

We can see that 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, and, except for (f), is in no small part due to the further coarsening of the discrete distribution.

IV. 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.

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 $\frac{1-2sp}{\sqrt{nsp(1-sp)}}$. The skewness of a gaussian distribution equals zero so a binomial distribution fit well by a gaussian requires that the skewness be close to zero. For the best approximation, it follows then

$$1 - 2sp = 0 \quad \Rightarrow \quad sp = \frac{1}{2}.$$

Thus, skewness grows worse with quantal failures when $0 < p < \frac{1}{2}$, 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 affect seems smaller.

The kurtosis of a binomial distribution is equal to $3 - \frac{6}{n} + \frac{1}{np(1-p)}$ 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

$$p^2 - p + \frac{1}{6} = 0.$$

They are $p = 0.21$ 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 $s < 1$, then kurtosis moves away from the gaussian value. So we can conclude that the failure process hurts when $p < 0.21$, which is the usual biological condition (firing rates in neocortex are arguably around $p \approx 0.12$ per 2.5 msec).

V. 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.

A distribution function for such mixture has the following form

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

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

We again ran some simulations of such two-binomial mixtures using a few different parameter, and with a desired firing rate of 5%. The results were as follows:

a) $p_1 = 0.05$, $w = 0.7$, $p_2 = 0.25$, $Q_{\text{Best}} = 161$, $Q_{\text{Gauss}} = 156.623$

b) $p_1 = 0.05$, $w = 0.3$, $p_2 = 0.25$, $Q_{\text{Best}} = 373$, $Q_{\text{Gauss}} = 495.961$

c) $p_1 = 0.15$, $w = 0.7$, $p_2 = 0.25$, $Q_{\text{Best}} = 230$, $Q_{\text{Gauss}} = 241.772$

d) $p_1 = 0.15$, $w = 0.3$, $p_2 = 0.25$, $Q_{\text{Best}} = 374$, $Q_{\text{Gauss}} = 469.604$

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

Now assume that the failure process is operating, and consider mixtures with the same parameters as before but with $s = 0.3$. Then the quantiles are:

e) $sp_1 = 0.015$, $w = 0.7$, $sp_2 = 0.075$, $Q_{\text{Best}} = 51$, $Q_{\text{Gauss}} = 48.027$

f) $sp_1 = 0.015$, $w = 0.3$, $sp_2 = 0.075$, $Q_{\text{Best}} = 120$, $Q_{\text{Gauss}} = 149.352$

g) $sp_1 = 0.045$, $w = 0.7$, $sp_2 = 0.075$, $Q_{\text{Best}} = 75$, $Q_{\text{Gauss}} = 75.794$

h) $sp_1 = 0.045$, $w = 0.3$, $sp_2 = 0.075$, $Q_{\text{Best}} = 120$, $Q_{\text{Gauss}} = 142.363$

Note the improvement especially when the larger mean valued mixing distribution dominates, and even when it does not dominate, little harm is done.

The comparison between (c) and (g) is particularly impressive. As shown in the accompanying figure based on two thousand samplings, it seems quantal failures can convert an inherently bimodal distribution to an inherently unimodal distribution.

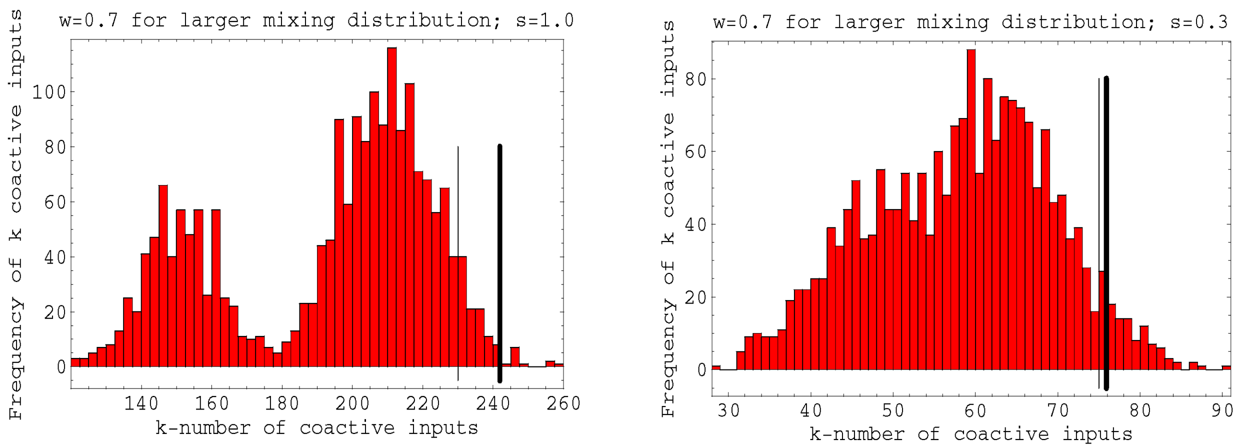

Figure 2. The effect of quantal failures on a two component mixture. Q_{Best}, for 5% average firing rate, is demarcated by a thin vertical line; Q_{Gauss} is demarcated by a thick vertical line. Note how these lines get closer when s = 0.3.

Thus for more complex input distributions, we conclude that the failure process can help the neuron get its average firing rate closer to the desired threshold.

VI. SKEWNESS AND KURTOSIS

Continuing with the excitation produced by these two component mixtures, 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. So, expanding the definitions of skewness and kurtosis

$$\text{skewness} = \frac{E[X - E[X]]^3}{\sigma^3}, \qquad \text{kurtosis} = \frac{E[X - E[X]]^4}{\sigma^4}$$

and using the above moments, we can calculate the skewness and the kurtosis for the mixture. The resulting equations are quite complicated so we skip to the results for the parameters of (a) through (h) on page 8:

a) skew = 0.877, kur = 1.795	e) skew = 0.895, kur = 1.895
b) skew = − 0.850, kur = 1.763	f) skew = − 0.782, kur = 1.764
c) skew = 0.848, kur=1.876	g) skew = 0.802, kur = 2.171
d) skew = − 0.804 kur=1.825	h) skew = − 0.620, kur = 1.962

Note that in all cases, except the skewness of (a) versus (e), quantal failures help to produce a closer approximation to a gaussian distribution. That is, in all the other cases, skewness moves towards zero when the failure process is added. The effect on k is even better, and Kurtosis is uniformly improved. Recalling that the kurtosis of a gaussian is three, we note that the failure process improves all four – rather different – examples.

In conclusion, we find that when n is large and the inputs are statistically independent, quantal failures do little harm to the hypothesized mechanism for a neuron to achieve its energy-efficient firing rate. More importantly where statistical independence does not hold, quantal failures will, in many or even most cases, improve the accuracy of this approximation.

REFERENCES

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

Gallagher RG Information Theory and Reliable Communication. John Wiley & Sons, Inc., New York, NY, 1968.

Katz, B. Nerve, Muscle, and Synapse. New York, NY: McGraw-Hill, Inc., 1966.

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

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

*Also published as part of a collection, In: Neural Codes and Distributed Representations. (L. Abbott and T. J. Sejnowski, Eds.), Cambridge, MA, MIT Press, pp. 105-117, 1999.

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

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

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