Synaptic Failures and a Gaussian Excitation Distribution

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. The energetic efficiency of an axon depends on its average firing rate [5]. We hypothesize that the failure process of quantal synaptic transmission moves a neuron's input excitation distribution closer to a gaussian, which then helps a neuron more precisely achieve an energetically desirable firing rate. If there are many statistically independent inputs per neuron, excitation is binomial and well approximated by a gaussian distribution. Quantal failures are unnecessary but are essentially harmless. Such statistical independence is, however, too simplistic. To reflect statistical dependence among the inputs, we consider mixture distributions. Generally, mixture distributions are a distributional class that can be far from gaussian even though the individual component distributions themselves are gaussian, or nearly so. Here we show that quantal synaptic failures can move the kurtosis and skewness of mixture distributions towards gaussian values.

1. Introduction

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 [6]. In order to quantify this last idea, we took advantage of assumptions which imply that the total 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 [4]). Here the mechanism proposed for setting the firing rate is simple: a neuron has kept track of the mean and the variance of the excitation induced by its inputs, and the threshold is implicitly set by some biophysical mechanism that is equivalent to the neuron assuming a gaussian distribution parameterized by this mean and variance.

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 large, because their states are finite, and because there are so many inputs being summed, we have every reason but one to invoke a central limit theorem. Specifically, the statistical dependence among the inputs could prevent the application of a c.l.t. to their sum. Here we define some conditions in which quantal failure improves the gaussian approximation by changing the third and fourth moments.

2. Failure process

In the hippocampus and the 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 [8], especially when one takes into account the spontaneous rates of neurons [1,7,]. The failure of quantal synaptic transmission is a random process [3], and because it loses information, it is counterintuitive when it exists under physiological conditions. The quantal failure process is seen to be equivalent to the so-called Z-channel of information theory well-studied (e.g., [2]) 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 it supposes that 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 binomially distributed but with parameters *n* and *sp*. (The asterisk notation will denote that the quantal failure process is operating.)

3. Gaussian approximation

Let us 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 with 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 8,000 random numbers for a variety of binomial parameterizations. 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 significantly different from 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. We come to a similar conclusion for kurtosis, 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. Moreover, to evaluate the effects of kurtosis, we set these two equal and are led 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 that 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 $Bin(n, p_1)$ and $\mathrm{Bin}(\mathrm{n}, p_2)$:

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

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

By expanding the definitions of skewness and kurtosis,

skewness = $E[K - E[K]]^3 / \sigma^3$, and kurtosis = $E[K - E[K]]^4 / \sigma^4$,

we can calculate the skewness and the kurtosis for the mixture where σ^2 is the variance of this mixture and $E[K]$ is the mean of 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 $p_2 = 0.06$ with the results:

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

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

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

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: $Q_{\text{Best}} = 106$, $Q_{\text{Gauss}} = 103.97$; skewness=0.37, kurtosis=2.98

(d) negative skewness, failures=0.8: $Q_{\text{Best}} = 110$, $Q_{\text{Gauss}} = 110.94$; 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, the failure process can convert an inherently bimodal mixture distribution to an inherently unimodal distribution. Furthermore, 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

By studying the derivative of skewness with regard to s = 1 − failure rate, we can predict the effect of failures. As we have reported previously [9],

Quantal failures can reduce skewness in two cases (each with a pair of conditions to satisfy) *but not in a third case*:

Case 1: $skewness > 0, \ and \ \dfrac{d(skewness)}{ds} > 0$; *equivalently*, *skewness* > 0, *and* 3 var > *mean.*

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. 1 ABOUT HERE.)

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

Moreover, skewness can always be driven to zero. That is, as $s \to 0, \ skewness > 0$ although at *s* = 1 *skewness* < 0. Figure 1 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.* Continuing with excitation produced by these two component mixtures, we also calculate the derivative of kurtosis with respect to *s*.

Kurtosis. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak.

Gaussian distributions have a kurtosis of 3 (irrespective of their mean or standard deviation). If a distribution`s kurtosis is greater than 3, it is said to be leptokurtic. If its kurtosis is less than 3, it is said to be platykurtic. Leptokurtosis is associated with distributions that are simultaneously “peaked” and have “fat tails”. Platykurtosis is associated with distributions that are simultaneously less peaked and have thinner tails.

The derivative of kurtosis with respect to *s* is:

$$\frac{d(kurtosis)}{ds} = \frac{d\mu_4}{ds} \cdot \frac{1}{var^2} - 2 \cdot \frac{dvar}{ds} \cdot \frac{\mu_4}{var^3}$$

where μ_4 is the fourth central moment of the mixture of two binomial distributions.

The derivative of kurtosis at *s=1* is

$$\frac{d(kurtosis)}{ds} = -6 \cdot skewness \cdot var^{-1/2} + \frac{4}{var} - \frac{mean}{var^2} - \frac{6 \cdot mean}{var} + \frac{2 \cdot mean \cdot kurtosis}{var}$$

Setting the derivative of kurtosis equals zero and multiplying by $\frac{var}{2 \cdot mean}$ gives

$$0 = \frac{-6 \cdot skewness \cdot (var)^{1/2}}{2 \cdot mean} + \frac{2}{mean} - \frac{1}{2 \cdot var} - 3 + kurtosis$$

It is to easy to observe that the terms $\frac{2}{mean}$ and $\frac{1}{2 \cdot var}$ are essentially equal to zero for the neurons we are interested in (*mean* and *var* are converging to infinity). The skewness is relatively small and if it is equal to zero then the sign of the derivative is the sign of the zero-centered kurtosis, which means quantal failures always help. Thus, in the case of kurtosis, the quantal failures process can move the excitation distribution closer to Gaussian in all cases that we consider reasonable. Figures 2 and 3 each show an example of an interaction between s and kurtosis.

(INSERT FIG. 2 AND 3 ABOUT HERE.)

7. Conclusions

In conclusion, when *n* is large and the inputs are statistically independent, quantal failures do little harm to gaussian-based thresholding so setting threshold to achieve an energy-efficient firing rate is easy. More importantly, where statistical independence does not hold, quantal failures will often improve the gaussian approximation.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] 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.

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

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

[4] W.B Levy, Contrasting rules for synaptogenesis, modification of existing synapses, and synaptic removal as a function of neuronal computation, Neurocomputing (2004) in press.

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

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

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

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

[9] J. Tyrcha and W.B Levy, Another Contribution by Synaptic Failures to Energy Efficient Processing By Neurons, Neurocomputing (2004) in press.

Joanna Tyrcha earned a BA in mathematical statistics from Main School of Planning and Statistics, Warsaw, Poland, and Ph.D. in mathematics from Jagiellonian University, Cracow, Poland. She was a research assistant at Institute of Computer Science of Polish Academy of Sciences from 1980 until 1992 at which point he joined the department of Numerical Analysis and Computing Science of Royal Institute of Technology, Stockholm, Sweden. She started to work at Institute of Mathematics, University of Stockholm, Stockholm, Sweden in 1996 where she is currently an associate professor in mathematical statistics.

William B Levy earned a B.A. in Psychology from Princeton University and a Ph.D. in Psychobiology from the University of California, Irvine. He served as Assistant Professor of Psychology at the University of

California, Riverside from 1974 until 1979. In 1979 he joined the faculty at the University of Virginia where he is currently a professor in the departments of Neurological Surgery and of Psychology.

FIGURE LEGENDS

Figure 1. 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: $n = 8000$, $p_1 = 0.05$, $p_2 = 0.06$. Mixing proportions: Case 1, $w = 0.7$; Case 2, $w = 0.3$.

Figure 2 Improvement in kurtosis when quantal failures are less than one. Kurtosis – 3 equals zero when $s = 0.148$. Simulations as in Case 1 of Fig. 1: $n = 8000$, $p_1 = 0.05$, $p_2 = 0.06$, $w = 0.7$.

Figure 3 Another example of quantal failures moving kurtosis to zero. Kurtosis – 3 equals zero when $s = 0.050$. Simulations as in Case 2 of Fig. 1: n = 8000, p1 = 0.05, p2 = 0.06, w = 0.7.

Figure 1

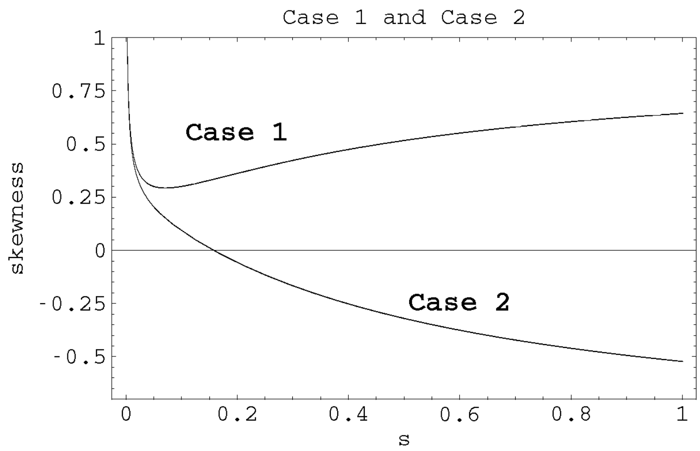

Figure 2

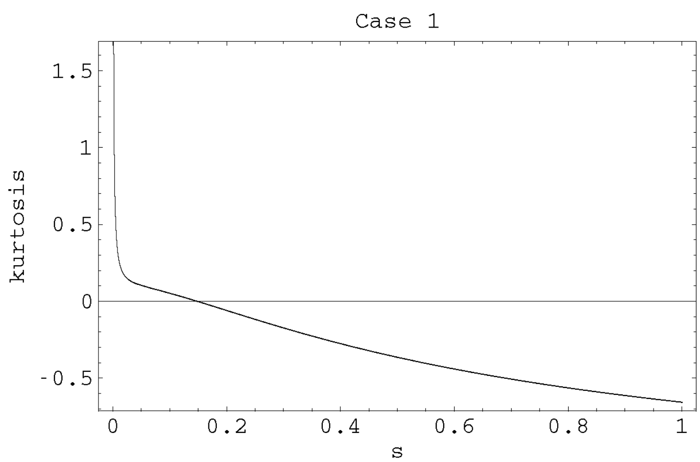

Figure 3