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, mixture

Abstract. If there are many statistically independent excitatory inputs to a neuron, its net excitation is binomially distributed, and this distribution is well-approximated by a Gaussian because there are so many inputs. Because there are so many inputs, quantal failures are essentially harmless. However, the presumption of statistical independence is 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 might, themselves, be Gaussian. Here we show when quantal synaptic failures can move the kurtosis of a mixture distribution towards the Gaussian value.

1. Introduction

It is often conjectured that neurons compute probabilities (e.g., [10]) or that they make decisions based on the logarithms of probability ratios (e.g., [6]). To implement such theories, it is necessary to assume something about the probability distribution that defines a neuron's excitation. More to the point, the neuron itself must implicitly assume something about its excitation distribution to perform such calculations. In such a situation, the more closely the true excitation distribution corresponds to the neuron's implicitly presumed distribution, then the more accurate are the neuron's decisions on average.

Perhaps the most widely studied model of dendritic excitation is the linearly scaled summation of McCulloch and Pitts [9], and this idea now receives experimental support [8]. When such a linearly scaled summation exists as the excitation distribution, then there are reasons to argue that the excitation distribution will tend towards the Gaussian (normal) distributional form particularly when the number of inputs, n, is large. In fact, for many neurons n is large (~10,000) relative to the standards of statistical sampling (where $n = 30$ is often invoked to justify a Gaussian approximation); thus it seems plausible that a central limit result might assert itself for postsynaptic excitation.

When a central limit result is in force, a neuron's task is simplified, and issues of local statistical consistency can be avoided. That is, to parameterize its excitation distribution it need only monitor (or have some developmental process guarantee values of) the mean and variance of excitation [3,7]. Then, it can easily implement distributionally-based decisions. For example energy-efficient use of the axon [4,5] demands a certain spike rate (or equivalently, firing probability in spikes per computational interval, see [4]), and this probability can be achieved by knowledge of the excitation distribution: specifically, setting threshold is just a matter of using the appropriate quantile of the parameterized distribution. But we must be careful.

In addition to a large number of summed inputs, slightly more is needed for a central limit theorem: When summing a finite number of finitely valued variables with some type of normalization, a central limit theorem seems inevitable unless the input is a mixture distribution. Unfortunately, the world is complicated and mixture distributions seem to abound (e.g., hidden Markov models are a very rich class of mixture models, and it seems that such models are necessary for a variety of forms of pattern recognition). Thus mixture distributions are something that we can expect a neuron to encounter.

Because – for a single postsynaptic neuron – quantal synaptic failures are apparently independent across inputs, it seems natural to argue that failures counteract the undesirable effect of a mixture on a central limit result. That is, the statistical dependence inherent in the existence of a mixture distribution might be undone by the randomness of the failure process across inputs. To say this another way, if synaptic failures can drive each of the statistical moments to its Gaussian value, then the resulting distribution is Gaussian. Here we continue our quantification of this idea by examining two moments of the excitation distribution when a quantal failure process is added to an excitation distribution that is a mixture. In addition to moving skewness toward the Gaussian value [14], there are conditions when failures move kurtosis towards its Gaussian value. Here we quantify these conditions.

2. Failure process

In forebrain cortical systems, excitatory synaptic connections dominate and are remarkably unreliable. Usually such a synapse transmits, at most, a single standardized package called a quantum [12]. 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 [13], especially when one takes into account the average firing rates of neurons [1,11]. This failure of synaptic transmission appears to be a random process [2]. As a stochastic process operating on input $X_i \in \{0,1\}$, $X_i = 0$ is unchanged, but $X_i = 1$ can change 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. Because it loses information (e.g., [5]), the failure process is counterintuitive when it exists under physiological conditions.

3. Gaussian approximation

Here is an example that uses a Gaussian approximation to set threshold. Given many independent equiprobable inputs, a Gaussian distribution well-approximates a binomial distribution. For a binomial distribution with parameters *n* (number of inputs) and *p* (probability of an input firing), the expected value of excitation equals *np* and variance equals *np(1-p)*. Then for threshold θ and postsynaptic excitation $\sum_{i=1}^{n} X_i$, the average firing of the postsynaptic neuron is

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

where Φ is the standardized, cumulative normal distribution function. If the desired average firing rate is, e.g., 5%, threshold is set as $\theta = 1.6449 \cdot \sqrt{np(1-p)} + np$, and this threshold produces the appropriate firing rate. To see how the failure process affects such a Gaussian-based threshold see [9].

Denoting σ^2 as the variance of a scalar random variable Z and E [ ] the expectation operator, in what follows it will be useful to recall:

Skewness is defined as $E\left[\left(Z - E[Z]\right)^3\right] \big/ \sigma^3$; and Kurtosis is defined as $E\left[\left(Z - E[Z]\right)^4\right] \big/ \sigma^4$.

4. Effect of quantal failure process on skewness and kurtosis

As we pointed out previously [14], because the skewness and kurtosis of a Gaussian are known, we can see exactly how s ($= 1 -$ failure rate) changes these characterizations of a binomial distribution relative to its Gaussian approximation. The skewness of a binomial distribution $Bin(n, p)$ equals $\frac{1-2p}{\sqrt{np(1-p)}}$ and, if the failure process operates, it is $\frac{1-2sp}{\sqrt{nsp(1-sp)}}$. The kurtosis of a binomial distribution is equal to $3 - 6/n + 1/(np(1-p))$ and if the failure process operates, it is then $3 - 6/n + 1/(nsp(1-sp))$. Because the skewness of a Gaussian equals zero and its kurtosis equals 3, a binomial distribution is better fit by a Gaussian when the skewness is closer to zero and kurtosis closer to 3. However, when we consider biological firing rates for which p is small (certainly p is less than 0.5 and usually p is less than .21), quantal failures do not move binomial excitation towards the Gaussian values. Failures move the excitation distribution away from the Gaussian values, and this is true when the third moment is considered and similarly for kurtosis although the effect seems smaller (see [14]). Fortunately, such changes are very small (see [14]).

5. Mixture inputs

Binomial distributions are not very appealing as models of a neuron's input excitation because the statistical independence between individual inputs implied by a binomial distribution is unrealistic. A more interesting model of the input statistics assumes that inputs arise from a mixture distribution. Therefore, let's now examine the effect of quantal synaptic failures on such a model, and for simplicity, suppose the mixture consists of two binomial distributions. The distribution function of such a mixture is

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

where $0 < w < 1$ is the so-called mixing parameter. By expanding the definitions of skewness (skew) and kurtosis (kurt), we can calculate the skewness and the kurtosis for such a mixture. Here we simulate three such 2-component mixtures with the number of inputs n = 8000, a desired firing rate 0.05, and with p_1=0.05 while p_2=0.06, and where σ^2 denotes the variance of this mixture and $E[\Sigma X_i]$ its mean.

Case 1, w=0.7, (skewness > 0 and kurtosis < 3):

A. no failures: $E\left[\sum X_i\right]$ = 424.119, σ^2 = 1714.25 ; Q_{Best} = 500; Q_{Gauss} = 492.22; skew = 0.63, kurt = 2.33

B. failure=0.8: $E\left[\sum X_i\right] = 84.71$, $\sigma^2 = 137.08$; $Q_{Best} = 106$; $Q_{Gauss} = 103.97$; skew = 0.37, kurt = 2.98

Case 2, w=0.3, (skewness < 0 and kurtosis < 3):

C. no failures: $E\left[\sum X_i\right] = 456.09$, $\sigma^2 = 1771.99$; $Q_{Best} = 511$; $Q_{Gauss} = 525.33$; skew = – 0.52, kurt = 2.23

D. failure=0.8: $E\left[\sum X_i\right] = 91.15$, $\sigma^2 = 144.75$; $Q_{Best} = 110$; $Q_{Gauss} = 110.94$; skew = – 0.06, kurt = 2.82

Case 3, w=0.8, (skewness > 0 and kurtosis > 3):

E. no failures: $E\left[\sum X_i\right] = 416.161$, $\sigma^2 = 1384.476$; $Q_{Best} = 494$; $Q_{Gauss} = 477.37$; skew = 0.95, kurt = 3.23

F. failure=0.8: $E\left[\sum X_i\right] = 83.111$, $\sigma^2 = 121.816$; $Q_{Best} = 103$; $Q_{Gauss} = 101.27$; skew = 0.44, kurt = 3.28

Histograms A & C of Fig. 1 show the problem encountered when using the Gaussian approximation with mixture distributions (for Case 1 see [14]; Case 2 also appears in [14]). The Gaussian-based thresholds (Q_{Gauss}, thick line) are far from the best thresholds (Q_{Best}, thin line) when there are no failures. In the negatively skewed Case 2 (Fig. 1A), the threshold is too high and the neuron will fire below its desired rate. In the positively skewed Case 3 (Fig 1C), the threshold is too low and the neuron will fire above its desired rate. However failures improve the situation. In Fig. 1, B & D, when the failure process operates (s = 0.2), Q_{Gauss} moves close to Q_{Best}. In Case 2 the difference scaled by the standard deviation $\left(\left|Q_{\text{Best}} - Q_{\text{Gauss}}\right|/\sigma\right)$ goes from 0.34 to 0.08 while in Case 3 the scaled differences are 0.46 and 0.16. Thus, firing rate is closer to the desired value for both a positive and a negative skew example.

In addition to Q_{Gauss} and Q_{Best} moving closer together, the failure process can convert a bimodal mixture distribution to a unimodal distribution. Furthermore, quantal failures are producing a closer approximation to a Gaussian distribution in the sense that skewness moves towards zero and kurtosis moves closer to three.

6. Analyses

By studying the derivative of skewness with regard to s = 1 – failure rate, we predicted [14] the effect of failures. Quantal failures can reduce skewness in two cases (each with a pair of conditions to satisfy) but not in a third case:

Case 1: whenever $skewness > 0 \; and \; \dfrac{d\left(skewness\right)}{ds} > 0, \; or \; when$

$$skewness > 0 \; and \; 3\,\sigma^2 > \mu$$

quantal failures can reduce skewness but it cannot be driven to zero.

Case 2: whenever $skewness < 0 \; and \; \dfrac{d\left(skewness\right)}{ds} < 0, \; or \; when$

$$skewness < 0 \; and \; 3 \cdot \sigma^2 < \mu$$

quantal failures can reduce skewness; moreover, skewness can always be driven to zero.

Case 3: *There is no reduction in absolute value of skewness if the conditions of Case 1 or Case 2 fail to hold.*

(This Case 3 is not related to the Case 3 in section 5 above.)

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).

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

$$\frac{d(kurtosis)}{ds} = \frac{d\mu_4}{ds} \cdot \frac{1}{\sigma^4} - 2 \cdot \frac{d\sigma^2}{ds} \cdot \frac{\mu^4}{\sigma^6}$$

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

The derivative of kurtosis at $s=1$ is

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

Setting the derivative of kurtosis equal to zero and multiplying by $\dfrac{\sigma^2}{2 \cdot \mu}$ gives

$$0 = \frac{-3 \cdot skewness \cdot \sigma}{\mu} + \frac{2}{\mu} - \frac{1}{2 \cdot \sigma^2} - 3 + kurtosis$$

the terms $\frac{2}{\mu}$ and $\frac{1}{2 \cdot \sigma^2}$ are essentially equal to zero for the neurons we are interested in (μ and σ^2 are converging to infinity). The skewness will always be relatively small compared to kurtosis, and if it is equal to zero, then the sign of the derivative is the sign of the (kurtosis – 3). When these two signs agree, then quantal failures always help. Thus, in the case of kurtosis, the quantal failure process can move the excitation distribution closer to three in all cases that we consider reasonable. Figures 2 and 3 each show examples in which failures move kurtosis toward three. In general,

Thm: negative (kurtosis – 3) can be driven to zero because kurtosis ≥ skewness + 3 and failures can drive skewness to plus infinity.

We conjecture that, as very high failure rates drive skewness and kurtosis away from the Gaussian values (see Fig. 2 & 3), they are driving postsynaptic excitation to a Poisson distribution. Finally we note that our results, both for skewness and for kurtosis, generalize to *n*-component mixtures of binomial distributions.

7. Conclusions

Exact conditions have been derived that specify when the quantal failure process can improve certain Gaussian moment approximations. In regard to such Gaussian approximations, failures are most advantageously applied when skewness < 0 and when kurtosis < 3, but other inequalities of this pair can also be improved. Finally, because of these results, we conjecture that failures improve decision-making by neurons that are implicitly using Gaussian approximations.

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

[3] W.B Levy, Contrasting rules for synaptogenesis, modification of existing synapses, and synaptic removal as a function of neuronal computation, Neurocomputing 58-60 (2004) 343-350.

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

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

[6] W.B Levy, C.M. Colbert, and N.L Desmond, Elemental adaptive processes of neurons and synapses: A statistical/computational perspective. In: M.A. Gluck and D.E. Rumelhart, eds., Neuroscience and Connectionist Models (Lawrence Erlbaum Assoc., Inc., Hillsdale, NJ, 1990) 187-235.

[7] W.B Levy, H. Delic, and D.M. Adelsberger-Mangan, The statistical relationship between connectivity and neural activity in fractionally connected feed-forward networks, Biol. Cybern. 80 (1999) 131-139.

[8] J.C. Magee, Dendritic integration of excitatory synaptic input, Nature Rev. Nsci. 1 (2000) 181-190.

[9] W.S. McCulloch and W. Pitts, A logical calculus of the ideas immanent in neuron activity, Bull. Math. Biophys. 5 (1943) 114-133.

[10] Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev. 65, 1958, 386-408.

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

[12] C. F. Stevens and Y. Wang, Facilitation and depression at single central synapses, Neuron 14 (1995) 795-802.

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

[14] J. Tyrcha and W.B Levy, Another Contribution by Synaptic Failures to Energy Efficient Processing By Neurons, Neurocomputing 58-60 (2004) 59-66.

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 improve the approximate Gaussianity of positively and negatively skewed mixture distributions. Likewise, when firing threshold of a neuron is set by a quantile based on a Gaussian assumption and knowledge of the mean and the variance, the desired right-hand tail weight can be improved by quantal failures. Here we seek a firing rate of 5%. Note how the assumption-based threshold (thick line) moves towards the optimal threshold (thin line) when the quantal failure mechanism is included (compare A vs. B and C vs. D). In B and D, the Gaussian distribution with the corresponding mean and variance is indicated by the filled circles. See Section 5 of text for a description of Case 1 and Case 2.

Figure 2. Improvement in kurtosis by quantal failures when (kurtosis – 3) is negative. If kurtosis < 3, then some value of s < 1 can drive kurtosis to three. No such guarantee is possible if kurtosis > 3. Case 1 (positive skewness) and Case 2 (negative skewness) are described in Section 5 of the text.

Figure 3. Improvement in (kurtosis – 3) is small when kurtosis is greater than three. Note that in Case 3, skewness and (kurtosis – 3) each are near their best (i.e., smallest) value at about the same failure rate. See Section 5 of text for description of Case 3.

Figure 1

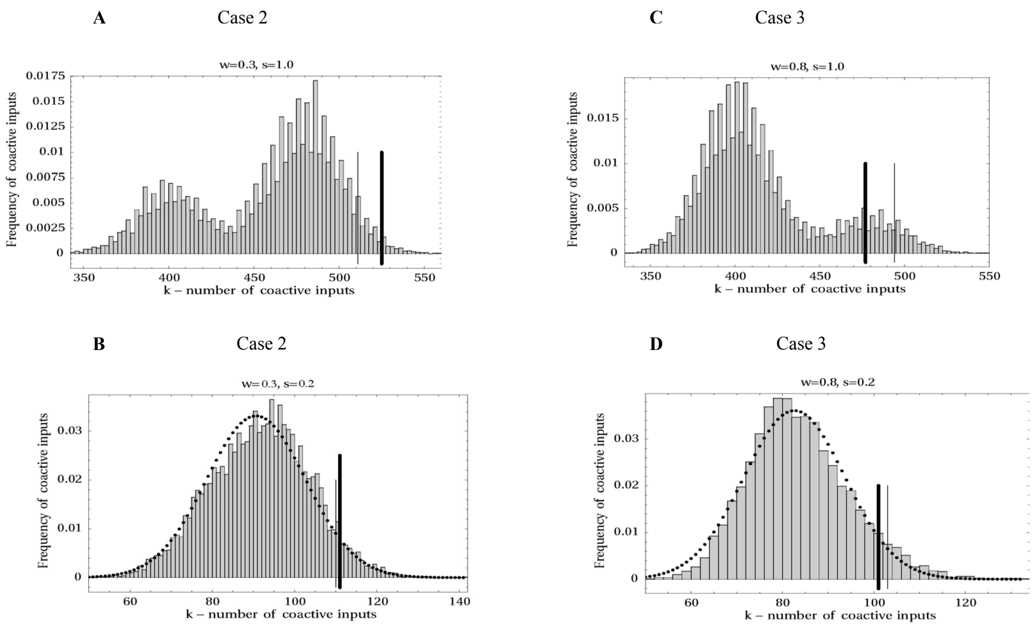

Figure 2

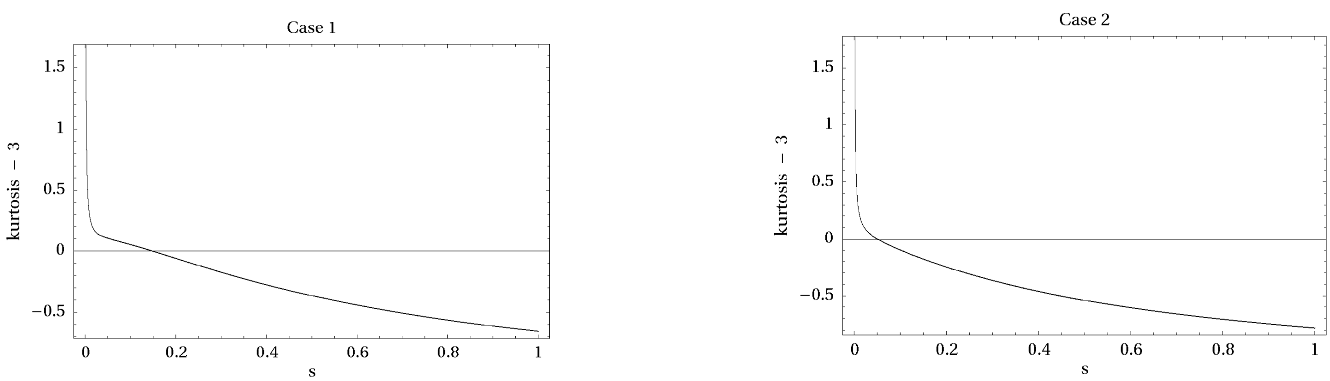

Figure 3