17 Balanced feedback networks: Noise, synaptic scaling, and linear response. Part 1

An interesting observation is that the subthreshold neuronal voltage in cortical neurons is very noisy. Far far noisier than one expects based on the thermal expectation of noise. Beyond membrane thermal noise, one might expect that the subthreshold potential would be noisy if there were relatively few synaptic inputs, consistent with the notion of a few strong inputs that one sees in cortical slice experiments. But the other possibility is that the input is so noisy because large excitatory inputs are offset by large inhibitory inputs, so that their mean value just about cancels but the variances, of course, add. The notion of large offsetting currents comes from the intracellular recording experiments in cat V1 from pioneering experiments in the laboratories of Ferster, of Fregnac, and of Douglas, with recent contributions from mouse by Isaacson and by Scanziani. In general, excitatory and inhibitory inputs are found to be both large and have the same tuning curves, so that their inputs act to balance each other. The gain from offsetting currents is that a transient increase in excitatory input, as may occur with a large burst of excitatory input, will rapidly depolarize the cell. So balanced networks trade noise for the speed gained with fast switching currents.

17.1 Variability for a single cell
17.1.1 Weak synaptic inputs

Let’s start with a warm up on the scale of noise in the input. First, some definitions, The input to cell $i$ from cell $j$ is $W_{ij}$ with $j = 1, 2, \ldots, N$, while the output of the neuron is taken as $V_i$ with $i = 1, 2, \ldots, N$ where $V = \frac{1}{2}(S + 1)$ is a Bernoulli variable with $V = 1$ if the cell spikes and $V = 0$ if it does not.

A Bernoulli probability distribution of the random variable $V$ can be thought of as a model for the set of possible outcomes of any single measurement whose outcomes is Boolean-valued. The Bernoulli distribution is a special case of the binomial distribution where a single trial is conducted, i.e., $N = 1$ for such a binomial distribution. Let’s define the probability that a cell is spiking as $m$, so that $V = 1$ with probability $m$ and $V = 0$ with probability $1 - m$.

The input to the $i-th$ neuron, denoted as in the past by $\mu_i(t)$, is:

$$\mu_i(t) \equiv \sum_{j=1}^{N} W_{ij} V_j(t). \quad (17.1)$$
The standard thermodynamic scaling, so that total synaptic currents are bounded as the size of the system increases, is that each input has a strength of order $1/N$. For simplicity, let’s take all of the inputs to be equal, so

$$W_{ij} \rightarrow \frac{W_o}{N}. \quad (17.2)$$

The $W_o$ are of order 1 so the total over all $N$ inputs is of order 1, with

$$\mu_i(t) = W_o \frac{1}{N} \sum_{j=1}^{N} V_j(t) \quad (17.3)$$

$$= W_o \ m(t)$$

where $m(t)$ is the average across the network, i.e.,

$$m(t) \equiv \frac{1}{N} \sum_{j=1}^{N} V_j(t) \quad (17.4)$$

and is of order 1. Clearly, for constant connection strengths, the input to all neurons is equal so the population average is

$$\mu_i(t) \equiv \frac{1}{N} \sum_{j=1}^{N} \mu_j(t) \quad (17.5)$$

$$\equiv \frac{1}{N} \sum_{j=1}^{N} W_o m(t)$$

$$= W_o \ m(t) \ \forall \ i$$

and the time average is

$$\langle \mu \rangle \equiv \frac{1}{T} \int_{-T/2}^{T/2} dt \ \mu_i(t) \quad (17.6)$$

$$\equiv W_o \frac{1}{T} \int_{-T/2}^{T/2} dt \ m(t)$$

$$= W_o \ m$$

The variance across time is

$$\sigma_i^2 = \langle (\mu_i(t) - \langle \mu \rangle)^2 \rangle \quad (17.7)$$

$$= [\langle \mu_i^2 \rangle - \langle \mu \rangle^2]$$

$$= \langle \mu_i^2 \rangle - W_o^2 \ m^2.$$
outputs are zero, i.e.,
\[
\langle \mu_i^2 \rangle = \langle \mu_i \mu_i \rangle
\] (17.8)
\[
= \left( \frac{W_o^2}{N^2} \sum_{j=1}^{N} \sum_{k=1}^{N} V_j(t) V_k(t) \right)
\]
\[
= W_o^2 \left( \frac{1}{N^2} \sum_{j=1}^{N} V_j^2(t) + \frac{1}{N^2} \sum_{j=1}^{N} \sum_{k \neq j} V_j(t) V_k(t) \right)
\]
\[
= W_o^2 \left( \frac{1}{N} \sum_{j=1}^{N} V_j(t) \right) + W_o^2 \left( \frac{N^2 - N}{N^2} \right) \left( \frac{1}{N} \sum_{j=1}^{N} V_j(t) \right)^2
\]
\[
= W_o^2 \langle m(t) \rangle + W_o^2 \left( 1 - \frac{1}{N} \right) \langle m^2(t) \rangle
\]
\[
= W_o^2 \frac{m - m^2}{N} + W_o^2 m^2
\] (17.9)

and thus
\[
\sigma_i^2 = \frac{W_o^2}{N} m(1 - m).
\] (17.10)

The variance for the population is the same, i.e.,
\[
\sigma^2 = \frac{W_o^2}{N} m(1 - m).
\] (17.11)

We see that for large networks the mean level drives the spiking and the variability goes to zero as 1/N, or equivalently the standard deviation goes to zero as 1/√N. As expected for a binomial variable, the variance is also zero when all neurons are active, i.e., \( m = 1 \), or quiescent, i.e., \( m = 0 \). Lastly, for a Poisson process, we get the slightly different answer of \( \sigma^2 = (W_o^2/N)m \) where \( m = \text{rate} \times \text{time interval} \).

Figure 1: Averaging over synapses decreases the RMS noise
17.1.2 Strong synaptic inputs

How can we have a network with high noise? Let’s recall the issue of networks with a small fraction of strong connections. The challenge is to recast the input so that the variance does not diminish to zero as a function of the number of input neurons. This is where the idea of balance comes into play.

1. We need the input to be the sums of two terms, one excitatory and one inhibitory.

2. We need the total current from these two terms to cancel, i.e., be equal and opposite in sign, to first order. The time-dependent variation in the firing rate of a neuron will reflect variations in the balance of the inputs.

3. We need a small fraction of active inputs, defined as \( K \), where \( 1 << K << N \).

4. With a small number of inputs, the total variance, which is the sum of variances of the excitatory and inhibitory terms, can be high.

Figure 2: Variability versus extent of excitation/inhibition balance. From Denuve and Machens, Nat Neuro 2016

The input to the \( i \)th neuron is now the sum of outputs from excitatory cells, i.e., the \( V^E_i(t) \), and inhibitory cells, i.e., the \( V^I_i(t) \). Thus

\[
\mu_i(t) = \mu^E_i(t) + \mu^I_i(t) = \sum_{j=1}^{K} W^E_{ij} V^E_j(t) + \sum_{j=1}^{K} W^I_{ij} V^I_j
\]

Let \( W^E_{ij} \) be an excitatory input and \( W^I_{ij} \) be an inhibitory input, simplified as above but now scaled to be large, where large is defined as order \( \frac{1}{\sqrt{K}} \) rather than order \( \frac{1}{K} \). Thus

\[
W^E_{ij} \to \frac{W^E_i}{\sqrt{K}} \quad \text{and} \quad W^I_{ij} \to -\frac{W^I_i}{\sqrt{K}}
\]
where the order parameters for excitation and inhibition are defined by

\[ m^E(t) \equiv \frac{1}{K} \sum_{j=1}^{K} V^E_j(t) \quad \text{and} \quad m^I(t) \equiv \frac{1}{K} \sum_{j=1}^{K} V^I_j(t) \] (17.15)

and we have assumed without loss of generality that the same number of excitatory and inhibitory inputs. The input is large if the excitatory and inhibitory terms do not cancel balance to within a factor of \(1/\sqrt{K}\). The variance, following the derivation for the single input case, is

\[
\sigma^2 = \frac{1}{K} \sum_{i=1}^{K} \left\langle \left( \mu_i^E(t) - \langle \mu^E \rangle \right)^2 \right\rangle + \frac{1}{K} \sum_{i=1}^{K} \left\langle \left( \mu_i^I(t) - \langle \mu^I \rangle \right)^2 \right\rangle \] (17.16)

\[ = \frac{(\sqrt{K} W_o^E)^2}{K} m^E (1 - m^E) + \frac{(\sqrt{K} W_o^I)^2}{K} m^I (1 - m^I) \]

\[ = (W_o^E)^2 m^E (1 - m^E) + (W_o^I)^2 m^I (1 - m^I). \]

The important point is that there is no decrement as \(K \to \infty\). Further, the variance remains nonzero for the special case of \(W_o^E m^E = W_o^I m^I\), where the network is in "perfect" balance.

### 17.1.3 Experimental evidence for \(\sqrt{k}\) scaling

It is fair to ask if there is evidence to support this scaling, which would depend on a homeostatic mechanism. The data comes from networks in cell culture of different size. The data supports scaling of the synaptic inputs, i.e., the post synaptic potentials, as \(1/K^{0.6}\). This is close to the predicted value of \(1/\sqrt{K}\) for strong inputs, as opposed to \(1/K\) for weak input. Not bad!

### 17.1.4 Variance versus mean driven spiking

How do we interpret the mean and variance in terms of spike probability? We use the approximation of neuronal output as a Bernoulli, i.e., \(V = 1\) if the cell spikes and \(V = 0\) if it does not. In the absence of noise the transition for 0 to 1 is sharp at \(\mu = \theta\). How does the average probability of spiking smear when the variance is
non-zero? The simplest possibility is to assume a Gaussian amplitude distribution, as we did in the study of the capacity of the Hopfield model, so that

\[
m(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\theta}^{\infty} dx \ e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

\[
= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\frac{\mu-\theta}{\sqrt{2}\sigma}} dx \ e^{-x^2}
\]

\[
= \frac{1 + \text{erf} \left( \frac{\mu-\theta}{\sqrt{2}\sigma} \right)}{2}.
\]

When \( \sigma \) is small compared to \( \mu - \theta \), the transition from \( m(t) = 0 \) to \( m(t) = 1 \) is weakly smoothed, with

\[
m(t) \quad \sigma \ll \mu - \theta \quad 1 - \frac{\sigma}{\sqrt{2\pi}(\mu - \theta)} e^{\frac{(\mu - \theta)^2}{2\sigma^2}}.
\]

When \( \sigma \) is large compared to \( \mu - \theta \), the transition from \( m(t) = 0 \) to \( m(t) = 1 \) is completely smoothed with

\[
m(t) \quad \sigma \gg \mu - \theta \quad \frac{1}{2}.
\]

An interesting issue is to have a fixed input and vary the noise. We see, numerically, that the spike rate increases monotonically with increasing values of \( \sigma \) to a saturation value of \( m = 0.5 \). Most interestingly, there is a roughly linear region of increase for mean rates between \( m = 0.05 \) and \( m = 0.25 \).
Figure 4: Gaussian noise threshold model to estimate effect of noise in driving neuronal responses.

Figure 5: Mean versus noise driven spiking in spinal cord slice. From Petersen and Berg, eLIFE, 2016.

Figure 6: Mean versus noise driven spiking in brain slice. From Lundstrom, Higgs, Spain and Fairhall, Nature Neuroscience 2008.