Revised 9 February 2024

4 Variability is a fundamental aspect of neural responses

We previously learned that neurons use two voltage levels and at voltage dependent conductance, to shift between the two levels. We now re-examine this viewpoint in terms of synaptic transmission and noise immunity on the one hand, and the trade-off between noise and the speed of a networks response on the other hand.

4.1 Scale of Thermal Fluctuations

The fundamental voltage scale is the thermal scale, or

$$\frac{k_B T}{e} \approx 25\ mV. \tag{4.1}$$

We now consider the smallest scale, that of thermal noise, in driving intrinsic fluctuations in the membrane voltage. Ion flow across the membrane is defined by a net conductance, G, across the cell. One way to derive the equation for the thermal noise is to use the equipartition theorem to equate the fluctuating energy in the membrane to the thermal energy, i.e.,

$$\frac{1}{2} C \delta V^2 = \frac{1}{2} k_B T \tag{4.2}$$

This leads to a fluctuation in the potential (Figure 1) of size

$$\delta V = \sqrt{\frac{k_B T}{C}}. \tag{4.3}$$

A different derivation is given in Box 1. This noise has the same spectral power density at all frequencies. This is different that other sources of noise, like $1/f$ noise, that has origins in processes occurring of a variety of energy scales (Figure 2).

The capacitance is measure of geometry and electric susceptibility ϵ. It is given by $C = \epsilon_m$ (area/thickness), so that for a thin dielectric sphere of thickness L and radius a, $C = \epsilon_m \frac{4\pi a^2}{L}$. Thus

$$\begin{aligned} \delta V &= \sqrt{\left(\frac{k_B T}{e}\right)\left(\frac{L}{\epsilon_m}\right)\frac{e}{4\pi a^2}} \\ &= \frac{1}{2a}\sqrt{\left(\frac{k_B T}{e}\right)\left(\frac{e}{c_m}\right)\frac{1}{\pi}}. \end{aligned} \tag{4.4}$$

Figure 1: Thermal noise and the Gaussian amplitude distribution.

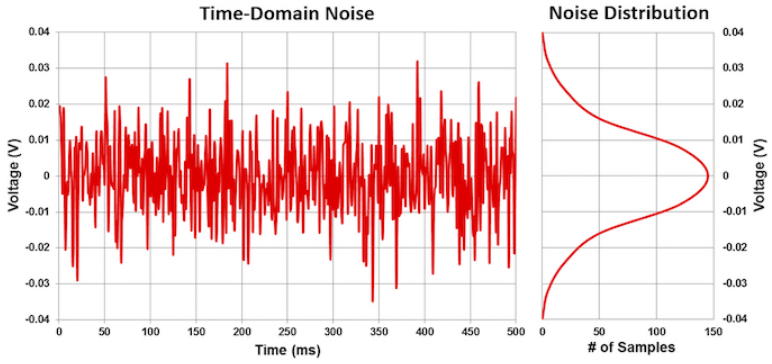
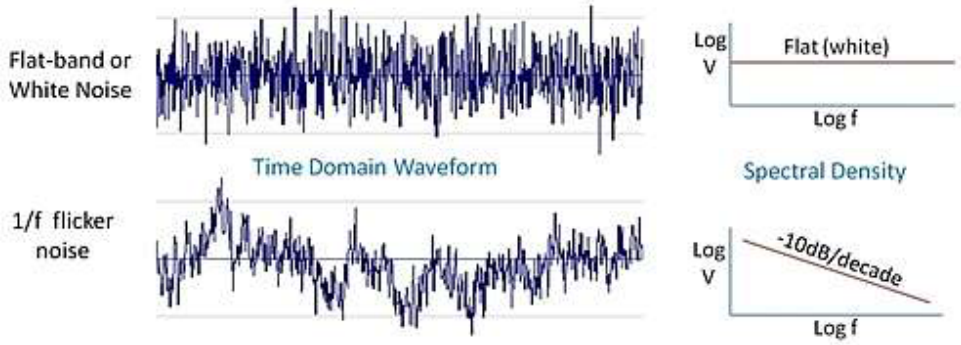


Figure 2: Thermal noise with a white - or flat - spectrum and flicker or Brownian noise - with a f^{-1} amplitude or f^{-2} spectrum.



For most all cells, the ratio $\frac{\epsilon_m}{L}$ is

$$\begin{aligned} c_m &\equiv \frac{\epsilon_m}{L} \\ &\approx 0.9x10^{-14}\frac{F}{\mu m^2} \end{aligned} \tag{4.5}$$

and

$$\frac{e}{c_m} = 1.8 \times 10^{-2}\frac{mV}{\mu m^2} \tag{4.6}$$

so that

$$\delta V \approx \frac{190\mu V}{a \text{ (in } \mu m)}. \tag{4.7}$$

For a cell of radius $a = 10\mu m$,

$$\delta V \approx 20\mu V. \tag{4.8}$$

Thus:

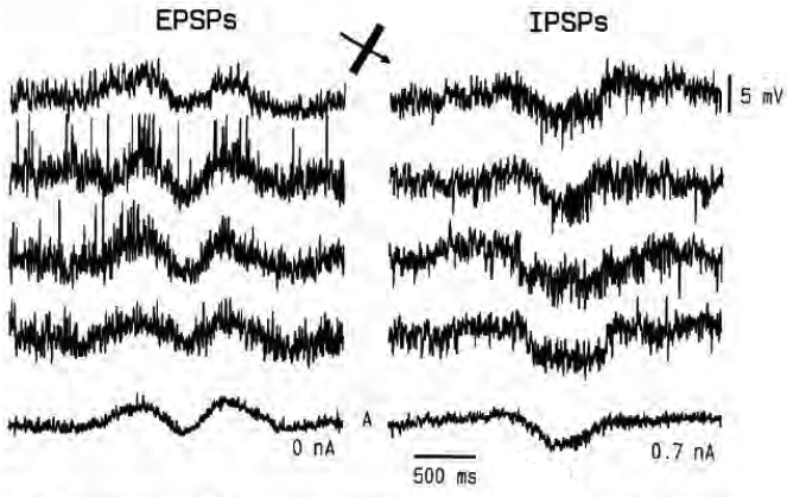
- The membrane noise level for cell somata is much less, by three orders of magnitude, then the thermal voltage $k_B T/e$.

- The membrane noise level for is less, by one order of magnitude, than the typical minimal synaptic input.

4.2 Variance versus mean driven spiking

Up to now we have considered driving neuronal by a change in the mean level of the input. We showed that a pulse of current will drive a neuron to fire and found the optimal shape for that pulse (Figure 8). In anticipation of a discussion of neuronal variability, consider how noise, or fluctuations in voltage, can drive a neuron to spike. Recall that noise has a zero mean value and is specified in terms of its range by the standard deviation or root-mean square value, denoted σ. We are concerned with noise on the scale of synaptic postsynaptic potentials, which sets the scale at 0.2 mV to 2 mV; the later value is similar to the transition from an inactive Na^+ current to a spike. In fact, intracellular measurements reveal an interesting fact. The postsynaptic potential is rapidly fluctuating with amplitudes of a few millivolts (Figure 3). This is surprising at first glance as neurons are believe to average over many inputs and thus one might imagine that the noise averages away; a Central Limit theorem arguments. But noise prevails, and as expected for a noisy subthreshold potential, the neuronal response to repeated presents of the same stimulus leads to a variable response (Figure 4).

Figure 3: The excitatory and inhibitory postsynaptic potentials for a neurons on primary visual cortex of cat. From Ferster 1988



4.2.1 Can noise alone can drive spiking?

Before we consider a mechanism for this noise, it is worth asking asking if noise alone can drive spiking? The answer is yes. When the average input to the neuron is well above threshold, the spiking is primarily driven by changes in the mean rate. But when the average input is held close to threshold, or just below threshold, fluctuations will drive the neuron to spike (Figure 5). In fact, the spike rate of the neurons can be a monotonic function of the standard deviation of the input (Figure 6).

Figure 4: Variability in spike rate with repeated presentation of the same visual random dot pattern. Data from monkey. From Shadlen and Newsome 1998

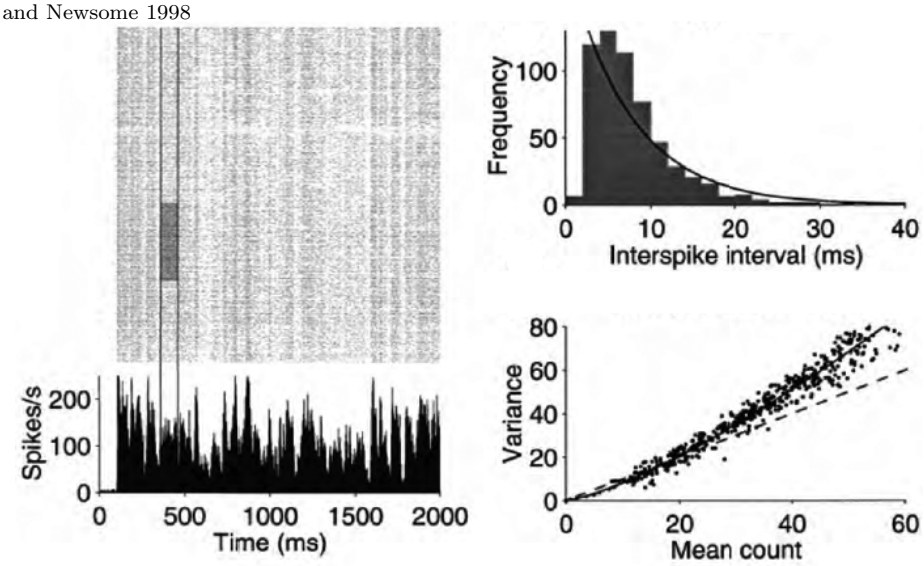
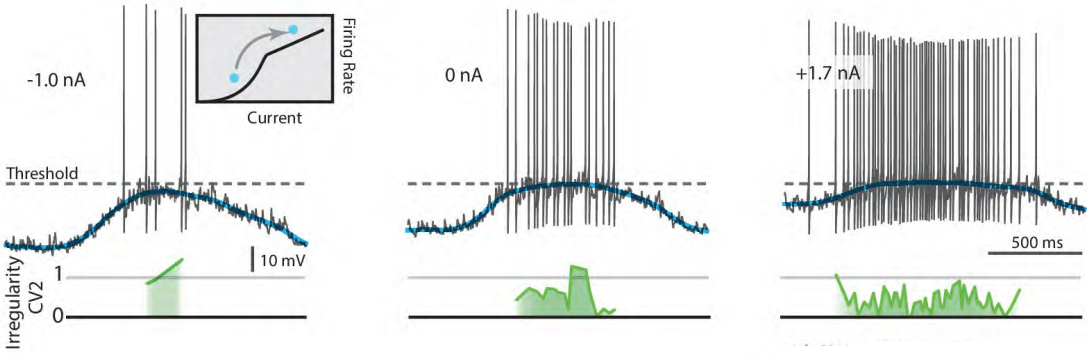


Figure 5: Mean versus noise driven spiking in spinal cord slice. From Petersen and Berg, eLIFE, 2016



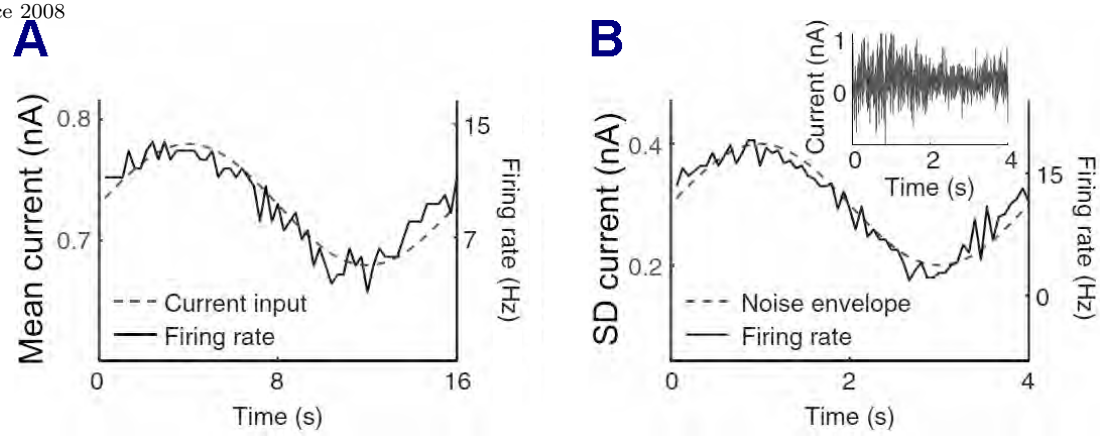
How do we interpret the mean and variance in terms of spike probability? We use the approximation of neuronal output as a Bernoulli, i.e., $V = 1$ if the cell spikes and $V = 0$ if it does not. In the absence of noise the transition for 0 to 1 is sharp at $\mu = \theta$. How does the average probability 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. We take $m(t)$ as the average output across the network, i.e.,

$$m(t) \equiv \frac{1}{N} \sum_{j=1}^{N} V_j(t) \tag{4.9}$$

so that

$$m(t) \approx \frac{1}{\sqrt{2\pi}\sigma} \int_{\theta}^{\infty} dx \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.10}$$

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



.

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\frac{\mu-\theta}{\sqrt{2}\sigma}} dx\ e^{-x^2}$$
$$= \frac{1 + \mathrm{erf}\left[\frac{\mu-\theta}{\sqrt{2}\ \sigma}\right]}{2}.$$

When σ is small compared to $\mu - \theta$, the transition from $m(t) = 0$ to $m(t) = 1$ is weakly smoothed (Figure 7), with

$$m(t) \quad \overrightarrow{\sigma \ll \mu - \theta} \quad 1 - \frac{\sigma}{\sqrt{2\pi}(\mu - \theta)} e^{-\frac{(\mu-\theta)^2}{2\sigma^2}}. \tag{4.11}$$

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

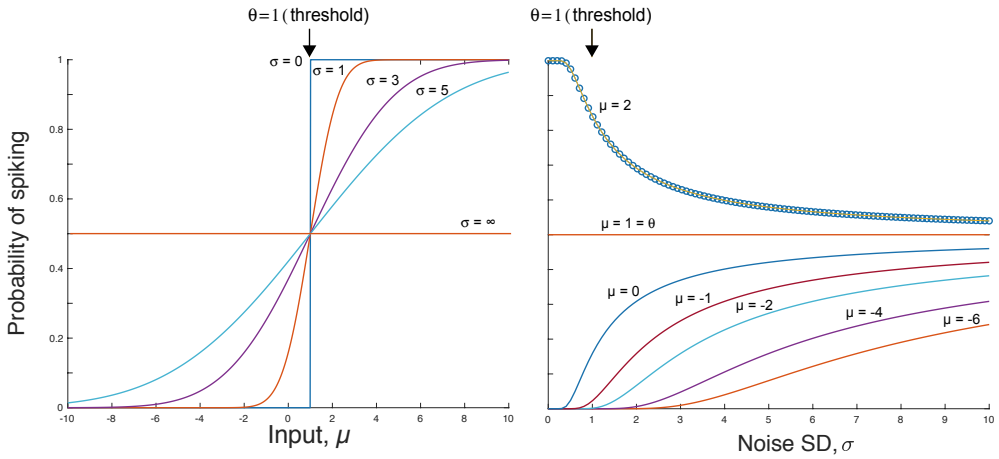
$$m(t) \quad \overrightarrow{\sigma \gg \mu - \theta} \quad \frac{1}{2}. \tag{4.12}$$

The interesting issue for us is to have a fixed input and vary the noise. We see, numerically, that the spike rate increases monotonically with increasing values of σ 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 $m = 0.25$.

4.3 The optimal input to drive spiking

We have considered that a step input leads to spiking and that noise can lead to spiking. But details of the ionic membrane currents should lead to a definitive input as the best driver of a neuron. While we have not explored such currents so far, the need to de-inactivate the inactivating component of the sodium channel suggests that a brief hyperpolarization before a depolarization is optimal. Indeed, using a noise-based correlation technique (Box 2) with a computer model of a Hodgkin Huxley cell, Blaise Aguera y Arcas and Adrienne Fairhall numerically determined the optimal current to drive a neuron (Figure 8). The combination of inhibitory and excitatory components to the input current suggests that this indicates the necessity of coordinating inhibitory and excitatory inputs

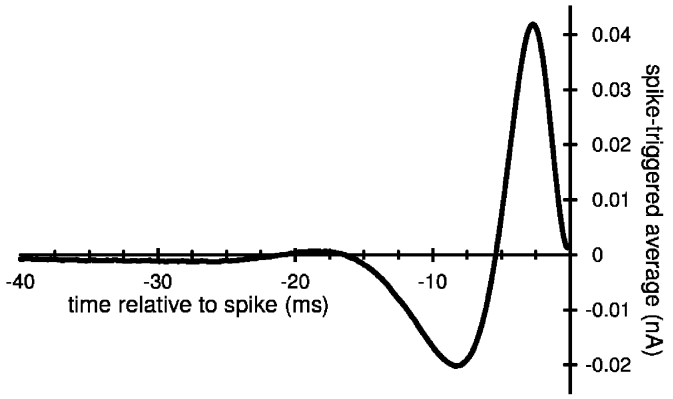
Figure 7: Gaussian noise threshold model to estimate effect of noise in driving neuronal responses



.

in brain circuits. Interestingly, the literature speaks of "feed-forward inhibition" and, as we shall see next, the tuning curves for inhibitory neurons in sensory and motor brain regions typically match those for excitatory neurons.

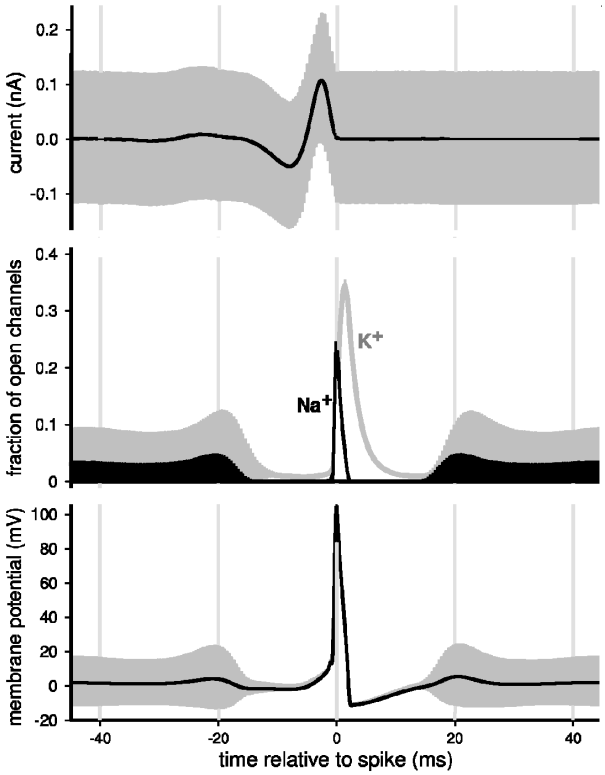
Figure 8: The optimal linear transfer function, i.e., current waveform, to elicit a spike. From Aguera y Arcas, Fairhall and Bialek (2003).



Another issue concerns the variability around the time of the spike. Consistent with an optimal input, Blaise Aguera y Arcas and Adrienne Fairhall observe that the variability

in membrane voltage is quenched at the time of a spike (Figure 9). This implies that the optimal input will reduce the jitter of spike timing.

Figure 9: Spike-triggered average of a Hodgkin-Huxley action potential with standard deviations for (top) the input current I, (middle) the fraction of open K^+ and Na^+ channels, and (bottom) the membrane voltage V, for the input current parameters mean $I_o = 0$ and spectral variance $S = 6.5 \times 10^{-4} nA^2 s$. From Aguera y Arcas, Fairhall and Bialek (2003).



4.4 Variability for a single cell

One might expect that the subthreshold potential would be noisy, if there were relatively few synaptic inputs. This is consistent with the notion of a few strong inputs that one sees in cortical slice experiments. Another possibility is that the subthreshold potential 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 (Figure 10). The notion of large offsetting currents comes from the intracellular recording experiments initially in anesthetized animals (Figure 3) and more recently in awake animals (Figure 21). In general, excitatory and inhibitory inputs are found to be both large and have the same sensory receptive fields or "tuning curves", so that their inputs act to balance each other, although this balance is not necessarily exact (Figure 12).

What is gained from this organization of offsetting currents? A transient increase in excitatory input, as may occur with a large burst of excitatory input, will rapidly depolarize the cell. So networks with balanced excitatory and inhibitory inputs, which

Figure 10: Balanced excitatory and inhibitory currents can lead to noisy input currents; calculated consequences of tight versus loose balance of excitatory and inhibitory currents. From Denuve and Machens 2016.

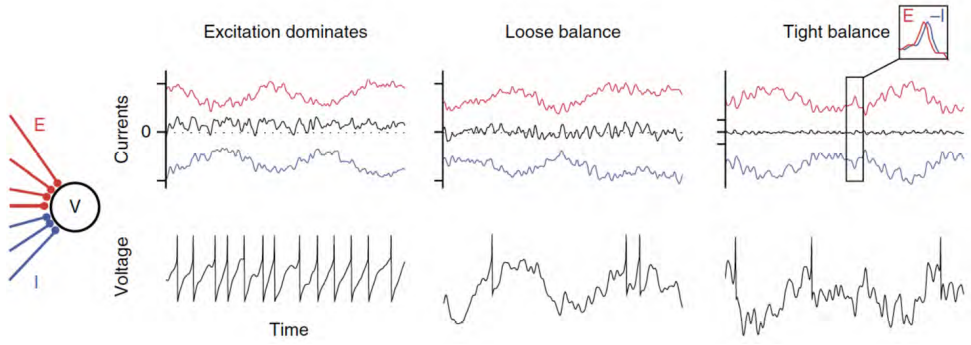


Figure 11: Balanced currents are observed in vivo in terms of balance of the gamma rhythm. From Atallah and Scanziani 2009.

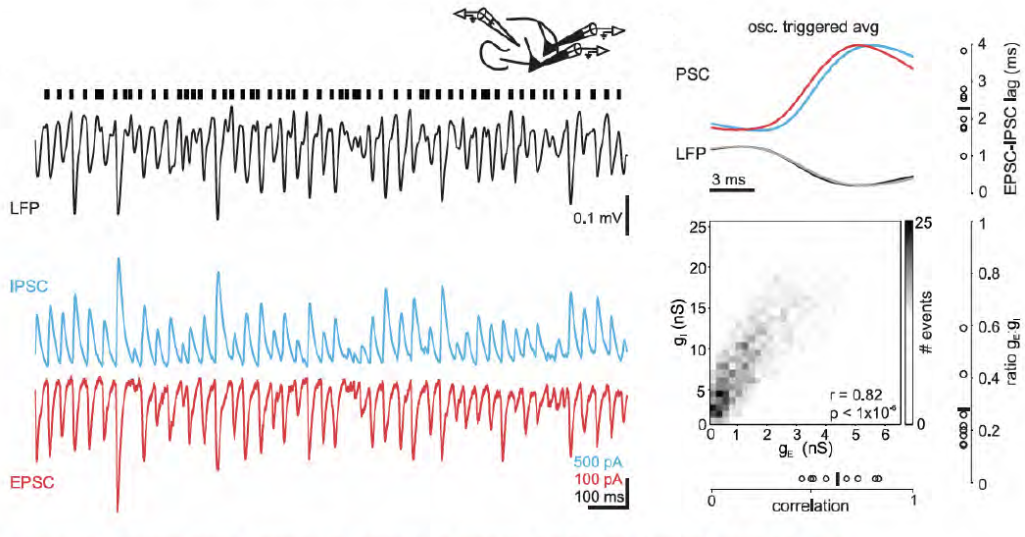
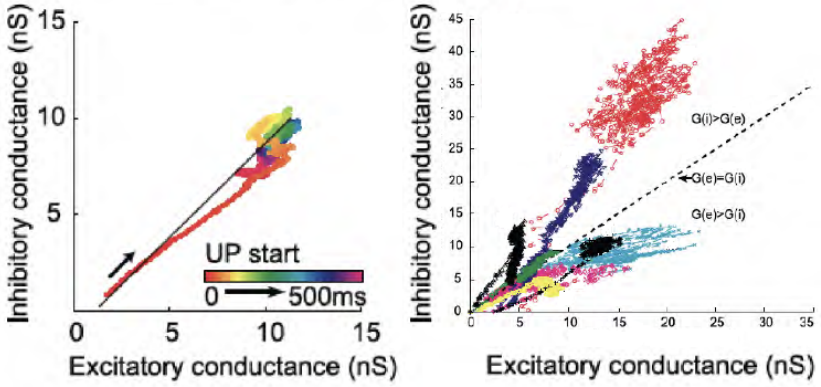


Figure 12: Balanced currents are proportional but do not necessarily exactly balance each other. Data from anesthetized mouse cortex. From Haider, Duque, Hasenstaub and McCormick 2006.



mean large conductances, are believed to trade noise from the balance for the speed gained from a large total leak conductance. We shall see!

4.4.1 Weak synaptic inputs

Let's start with a warm up on the scale of noise in the input. We use a rate model. 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 take as 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:

$$\mu_i(t) \equiv \sum_{j=1}^{N} W_{ij} V_j(t). \tag{4.13}$$

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

$$W_{ij} \to \frac{W_o}{N}. \tag{4.14}$$

The W_o are of order 1 in magnitude, so the sum over all N inputs is of order 1, with

$$\mu_i(t) = W_o \; m(t) \tag{4.15}$$

where $m(t)$ is the order parameter given by the average across the network, i..e.,

$$m(t) \equiv \frac{1}{N} \sum_{j=1}^{N} V_j(t) \tag{4.16}$$

so

$$\mu_i(t) \;=\; W_o m(t) \; \forall \; i \tag{4.17}$$

Clearly, for constant connection strengths, the input to all neurons is equal so the population average is

$$\mu(t) = W_o \; m(t) \tag{4.18}$$

and the time average is

$$\begin{aligned} \langle \mu \rangle \;&\equiv\; \frac{1}{T} \int_{-T/2}^{T/2} dt \; \mu(t) \\ &\equiv\; W_o \, \frac{1}{T} \int_{-T/2}^{T/2} dt \; m(t) \\ &=\; W_o \; m \end{aligned} \tag{4.19}$$

The variance across time is

$$\begin{aligned} \sigma^2 &= \left\langle (\mu_i(t) - \langle \mu \rangle)^2 \right\rangle \\ &= \left[\left\langle \mu^2 \right\rangle - \langle \mu \rangle^2 \right] \\ &= \left\langle \mu^2 \right\rangle \; - \; W_o^2 \; m^2. \end{aligned} \tag{4.20}$$

We evaluate the first term under the assumption that the correlations in the neuronal outputs are zero, i.e.,

$$\begin{aligned} \left\langle \mu^2 \right\rangle &= \left\langle \frac{W_o}{N} \sum_{j=1}^{N} V_j(t) \; \frac{W_o}{N} \sum_{k=1}^{N} V_k(t) \right\rangle \\ &= \left\langle \frac{W_o^2}{N^2} \sum_{j=1}^{N} \sum_{k=1}^{N} V_j(t) V_k(t) \right\rangle \\ &= W_o^2 \left\langle \frac{1}{N^2} \sum_{j=1}^{N} V_j^2(t) + \frac{1}{N^2} \sum_{j=1}^{N} \sum_{k \neq j}^{N} V_j(t) V_k(t) \right\rangle \\ &= \frac{W_o^2}{N} \left\langle \frac{1}{N} \sum_{j=1}^{N} V_j(t) \right\rangle + W_o^2 \left(\frac{N^2 - N}{N^2} \right) \left\langle \left(\frac{1}{N} \sum_{j=1}^{N} V_j(t) \right)^2 \right\rangle \\ &= \frac{W_o^2}{N} \langle m(t) \rangle + W_o^2 \left(1 - \frac{1}{N} \right) \left\langle m^2(t) \right\rangle \\ &= \frac{W_o^2}{N} (m - m^2) \; + \; W_o^2 m^2 \end{aligned} \tag{4.21}$$

(4.22)

and thus variance for the population is

$$\sigma^2 = \frac{W_o^2}{N} m(1 - m). \tag{4.23}$$

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 $1/\sqrt{N}$ (Figure 13). 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 $m =$ rate $\times$ time interval.

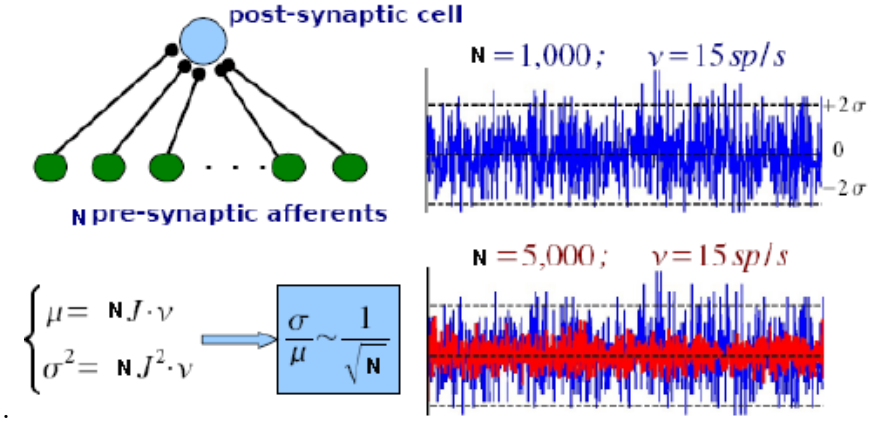
4.4.2 Strong synaptic inputs

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 balanced inhibition and excitation 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 term 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.

Figure 13: Averaging over synapses decreases the RMS noise



3. We need a small fraction of active inputs, defined as K, where $1 \ll K \ll N$.

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.

The input to the i-th neuron is now the sum of outputs from excitatory cells, i.e., the $V_i^E(t)$, and inhibitory cells, i.e., the $V_i^I(t)$. Thus

$$\begin{aligned} \mu_i(t) &= \mu^E(t) + \mu^I(t) \\ &= \sum_{j=1}^{K} W_{ij}^E V_j^E(t) + \sum_{j=1}^{K} W_{ij}^I V_j^I \end{aligned} \tag{4.24}$$

Let W_{ij}^E be an excitatory input and W_{ij}^I 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

$$W_{ij}^E \to \frac{W_o^E}{\sqrt{K}} \quad \text{and} \quad W_{ij}^I \to -\frac{W_o^I}{\sqrt{K}} \tag{4.25}$$

where we implicitly fix the sign of the inhibition. The mean input under the assumed scaling is

$$\begin{aligned} \mu_i(t) &= W_o^E \frac{1}{\sqrt{K}} \sum_{j=1}^{K} V_j^E(t) \; - \; W_o^I \frac{1}{\sqrt{K}} \sum_{j=1}^{K} V_j^I(t) \\ &= \sqrt{K} \left[W_o^E \frac{1}{K} \sum_{j=1}^{K} V_j^E(t) \; - \; W_o^I \frac{1}{K} \sum_{j=1}^{K} V_j^I(t) \right] \\ &= \sqrt{K} \left[W_o^E m^E(t) \; - \; W_o^I m^I(t) \right] \end{aligned} \tag{4.26}$$

where the order parameters for excitation and inhibition are defined by are defined by

$$m^E(t) \equiv \frac{1}{K} \sum_{j=1}^{K} V_j^E(t) \quad \text{and} \quad m^I(t) \equiv \frac{1}{K} \sum_{j=1}^{K} V_j^I(t) \tag{4.27}$$

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

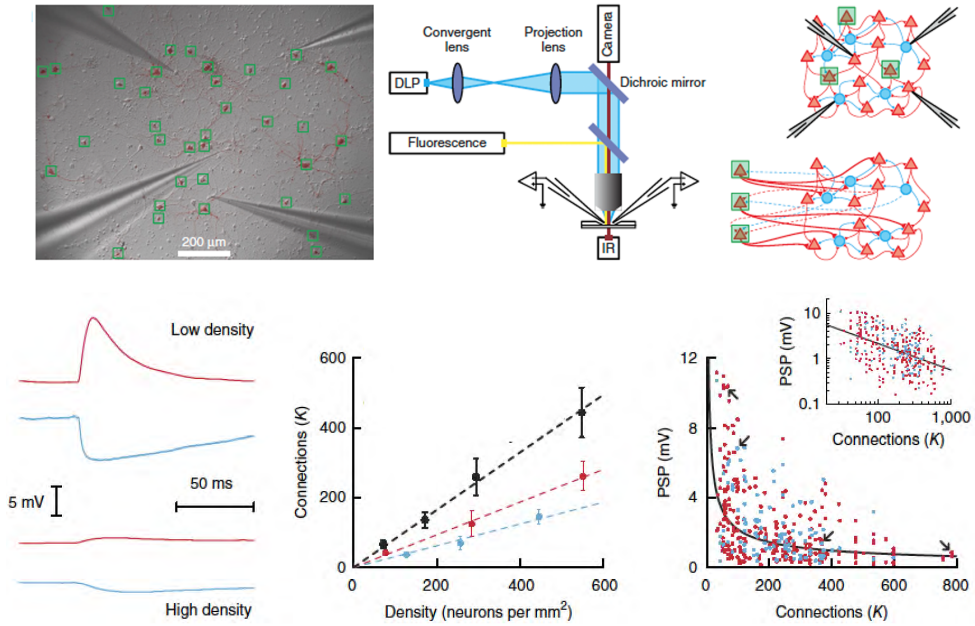
$$\begin{aligned}\sigma^2 &= \frac{1}{K}\sum_{i=1}^{K}\left\langle \left(\mu_i^E(t) - <\mu^E>\right)^2\right\rangle + \frac{1}{K}\sum_{i=1}^{K}\left\langle \left(\mu_i^I(t) - <\mu^I>\right)^2\right\rangle \qquad (4.28)\\ &= \frac{(\sqrt{K}W_o^E)^2}{K}\, m^E(1-m^E) + \frac{(\sqrt{K}W_o^I)^2}{K}\, m^I(1-m^I)\\ &= (W_o^E)^2\, m^E(1-m^E) + (W_o^I)^2\, m^I(1-m^I).\end{aligned}$$

The important point is that there is no decrement in the variance 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.

4.4.3 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 for maintenance. 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 $1/K^{0.6}$ (Figure 14). This is close to the predicted value of $1/\sqrt{K}$ for strong inputs, as opposed to $1/K$ for weak input. Not bad!

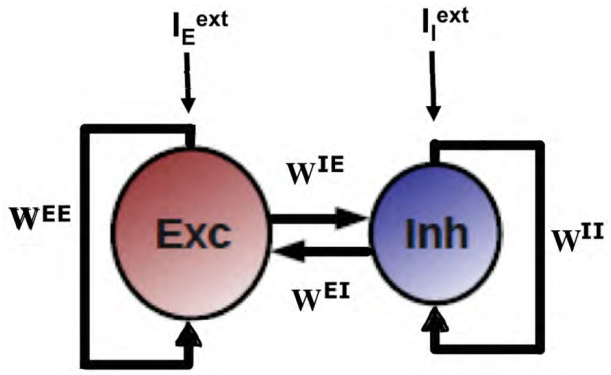
Figure 14: In vitro synaptic scaling preserves excitatory-inhibitory balance. From Barres and Reyes, 2016.



4.5 Circuit model

So far we have only address noise and scaling at the level of noise in individual cells. Now we analyze a network of neurons with balanced inputs (Figure 15). We consider the consequences of the choice of connections in a network on the ability to maintain the balanced state.

Figure 15: Feedback circuit model with two populations of neurons



.

Consider a network of a population of interconnected excitatory (E) and inhibitory (I) cells.The full equations are

$$\tau_E \frac{dV_i^E(t)}{dt} + V_i^E(t) = \left[\mu_i^E(t) - \theta_i^E\right]_+ \tag{4.29}$$

and

$$\tau_I \frac{dV_i^I(t)}{dt} + V_i^I(t) = \left[\mu_i^I(t) - \theta_i^I\right]_+, \tag{4.30}$$

where $[\cdots]_+$ is the Heavyside function, τ_E and τ_I are the cellular time constant, and the θ_i^E and θ_i^I are thresholds. The inputs are

$$\mu_i^E(t) = \mu_{ext}^E(t) + \sum_{j=1}^{K} W_{i,j}^{EE} V_j^E(t) + \sum_{j=1}^{K} W_{i,j}^{EI} V_j^I(t) \tag{4.31}$$

and

$$\mu_i^I(t) = \mu_{ext}^I(t) + \sum_{j=1}^{K} W_{i,j}^{II} V_j^I(t) + \sum_{j=1}^{K} W_{i,j}^{IE} V_j^E(t). \tag{4.32}$$

As in the case of the model cell, we will scale the synaptic inputs by $1/\sqrt{K}$, as opposed to $1/K$, i.e.,

$$W_{ij}^{EE} \to \frac{W^{EE}}{\sqrt{K}}; \quad W_{ij}^{II} \to -\frac{W^{II}}{\sqrt{K}}; \quad W_{ij}^{EI} \to -\frac{W^{EI}}{\sqrt{K}}; \quad W_{ij}^{IE} \to \frac{W^{IE}}{\sqrt{K}} \tag{4.33}$$

where we explicitly put in the negative signs of inhibition. As will soon be clear, we need to scale the external inputs by

$$\mu_{ext}^E(t) \to \sqrt{K}\ E\ m_{ext}(t) \quad \text{and} \quad \mu_{ext}^I(t) \to \sqrt{K}\ I\ m_{ext}(t) \tag{4.34}$$

where E and I are inputs of strength of $O(1)$. The dependence on a common term is a statement that excitatory and inhibitory neurons share the same tuning curve (Figure 3). All together, we have

$$\mu_i^E(t) = \sqrt{K}Em_{ext}(t) + \frac{W^{EE}}{\sqrt{K}}\sum_{j=1}^{K} V_j^E(t) - \frac{W^{EI}}{\sqrt{K}}\sum_{j=1}^{K} V_j^I(t) \tag{4.35}$$

and

$$\mu_i^I(t) = \sqrt{K}Im_{ext}(t) + \frac{W^{IE}}{\sqrt{K}}\sum_{j=1}^{K} V_j^E(t) - \frac{W^{II}}{\sqrt{K}}\sum_{j=1}^{K} V_j^I(t). \tag{4.36}$$

In terms of the order parameters,

$$\begin{aligned}\mu_E(t) &= \sqrt{K}Em_{ext}(t) + \sqrt{K}W^{EE}\frac{1}{K}\sum_{j=1}^{K} V_j^E(t) - \sqrt{K}W^{EI}\frac{1}{K}\sum_{j=1}^{K} V_j^I(t) \\ &= \sqrt{K}Em_{ext}(t) + \sqrt{K}W^{EE}m_E(t) - \sqrt{K}W^{EI}m_I(t) \\ &= \sqrt{K}\left[Em_{ext}(t) + W^{EE}m_E(t) - W^{EI}m_I(t)\right]\end{aligned} \tag{4.37}$$

and

$$\begin{aligned}\mu_I(t) &= \sqrt{K}Im_{ext}(t) + \sqrt{K}W^{IE}\frac{1}{K}\sum_{j=1}^{K} V_j^E(t) - \sqrt{K}W^{II}\frac{1}{K}\sum_{j=1}^{K} V_j^I(t) \\ &= \sqrt{K}\left[Im_{ext}(t) + W^{IE}m_E(t) - W^{II}m_I(t)\right].\end{aligned} \tag{4.38}$$

As $\sqrt{K} \to \infty$ the left hand side goes to zero and the equilibrium state will satisfy

$$0\left(\frac{1}{\sqrt{K}}\right) = Em_{ext}(t) + W^{EE}m_E(t) - W^{EI}m_I(t) \tag{4.39}$$

and

$$0\left(\frac{1}{\sqrt{K}}\right) = Im_{ext}(t) + W^{IE}m_E(t) - W^{II}m_I(t). \tag{4.40}$$

The implication of this equilibrium condition is that the average input remains finite as the fluctuations remain large (Figures 16 and 17). This is the balanced state.

4.6 The balanced state

Solving the above equations for m_E^o and m_I^o gives relations for the equilibrium activity of the excitatory and inhibitory cells in terms of the external drive:

$$m_E^0 = \frac{W^{II}E - W^{EI}I}{W^{EE}W^{II} - W^{EI}W^{IE}}m_{ext}. \tag{4.41}$$

and

$$m_I^0 = \frac{W^{IE}E - W^{EE}I}{W^{EE}W^{II} - W^{EI}W^{IE}}m_{ext}. \tag{4.42}$$

Recall that the equilibrium values of activity m_E^o and m_I^o must be both positive and bounded by 1. This constrains the values of the synaptic weights.

Figure 16: Balanced networks have emergent variability. From Shadlen and Newsome, 1994.

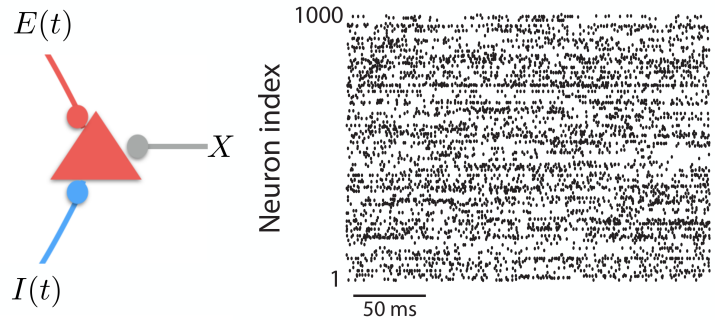
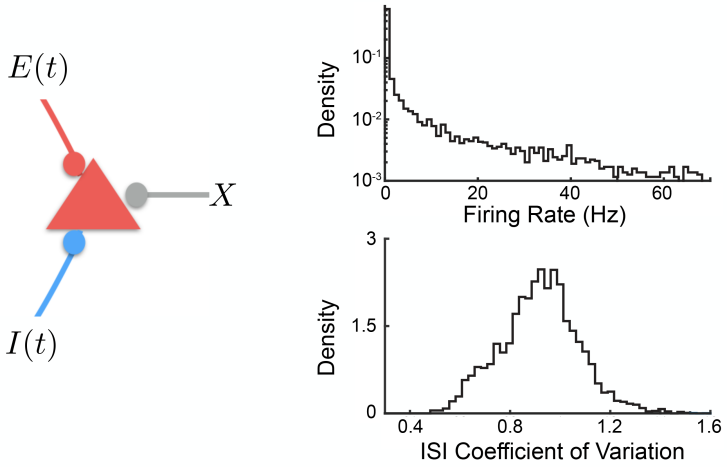


Figure 17: Statistics of have emergent variability. From Shadlen and Newsome, 1994.



4.6.1 Linear response

A seemingly paradoxical effect is that increasing the external inhibitory input, i.e., increasing I, will lead to a net decreased spiking of inhibitory cells and l will concurrently decrease both m_E and m_I (Figure 18). This is a feedback effect. Excitatory and inhibitory activity track each other until the excitatory cells are completely turned off; this behavior is seen across cortical regions (Figure 19).

A second issue is that rapid feedback prevents the occurrence of significant correlations. This depends of having faster inhibitory than excitatory synapses, as occurs for Gaba-A, but not Gaba-B (Figure 20).

4.6.2 Stability and response speed

We return to the full network equations and look at the variation around the equilibrium value of m_E and m_I. Taking the time constants, τ, conversion gains, β, and thresholds to be the same for the E and I populations, and denoting

$$\delta m_E(t) = m_E(t) - m_E^o \tag{4.43}$$

Figure 18: Model predictions of excitatory and inhibitory responses to inhibitory stimulation. (A) Schematic of model, showing connections between excitatory (E) and inhibitory (I) neuron populations. (B) Predictions for average neural responses with weak recurrent coupling (left) and strong coupling (right), when inhibitory cells are externally stimulated. (C) Schematic of experiment. Extracellular recordings made in visual (V1), primary somatosensory (SOMATO), and motor/premotor cortices (a: anterior, p: posterior, m: medial, l: lateral) while optogenetically stimulating inhibitory cells at the recording site in awake VGAT-ChR2 animals. From Sanzeni, Akitake, Goldbach, Leedy, Brunel and Histed 2020.

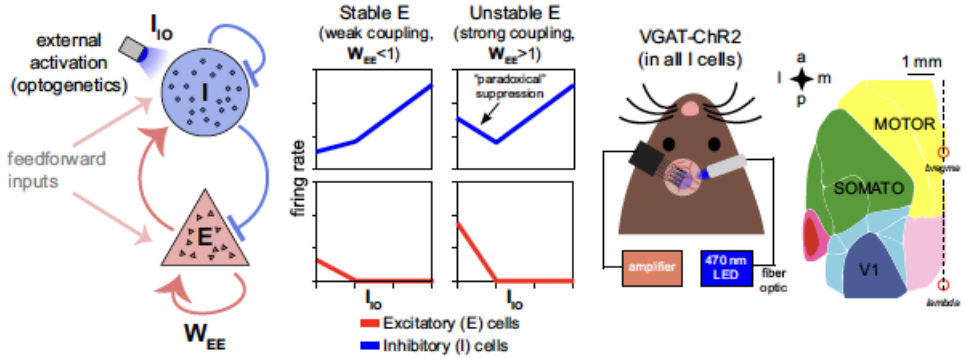


Figure 19: Inhibition stabilization across cortical areas. (left) Motor/premotor cortex recordings and motor cortex population firing rates for E and I units. Initial mean response of inhibitory cells is negative, showing paradoxical suppression. Mean rate is significantly reduced ($p_i 10^{-4}$, paired t-test, rate at 0 versus rate at L0). (right)Similar experiment but for recordings from somatosensory cortex; the mean I rate is significantly reduced ($p_i 10^{-7}$, paired t-test). From Sanzeni, Akitake, Goldbach, Leedy, Brunel and Histed 2020.

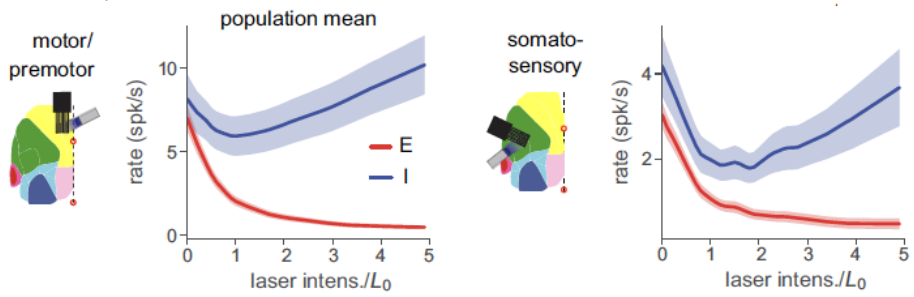
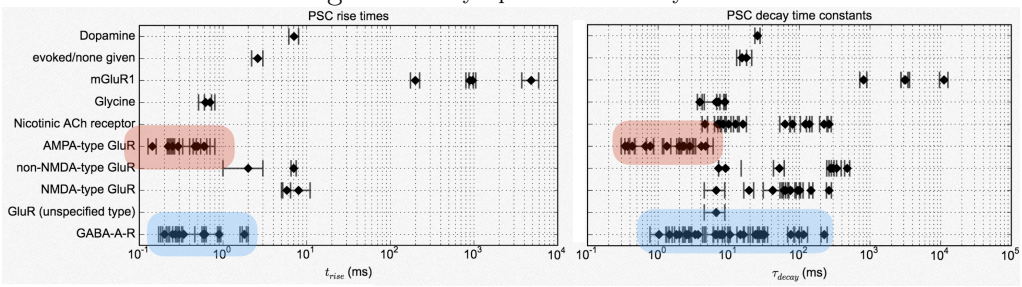


Figure 20: Synaptic rise and decay times



and

$$\delta m_I(t) = m_I(t) - m_I^o \tag{4.44}$$

leads to

$$\tau \frac{d\ \delta m_E(t)}{dt} + \delta m_E(t) = \left[\beta\sqrt{K}\left(W^{EE}\delta m_E(t) - W^{EI}\delta m_I(t)\right)\right]_+ \tag{4.45}$$

and

$$\tau \frac{d\ \delta m_I(t)}{dt} + \delta m_I(t) = \left[\beta\sqrt{K}\left(W^{IE}\delta m_E(t) - W^{II}\delta m_I(t)\right)\right]_+ . \tag{4.46}$$

When the neurons are active, this reduces to the linear equations

$$\tau \frac{d\ \delta m_E(t)}{dt} + \delta m_E(t) = \beta\sqrt{K}\left(W^{EE}\delta m_E(t) - W^{EI}\delta m_I(t)\right) \tag{4.47}$$

and

$$\tau \frac{d\ \delta m_I(t)}{dt} + \delta m_I(t) = \beta\sqrt{K}\left(W^{IE}\delta m_E(t) - W^{II}\delta m_I(t)\right). \tag{4.48}$$

These linear equations are solved by taking $\delta m_E(t) \propto e^{\lambda t}$, so that

$$(\lambda\tau + 1)\,\delta m_E(t) = \beta\sqrt{K}\left(W^{EE}\delta m_E(t) - W^{EI}\delta m_I(t)\right) \tag{4.49}$$

and

$$(\lambda\tau + 1)\,\delta m_I(t) = \beta\sqrt{K}\left(W^{IE}\delta m_E(t) - W^{II}\delta m_I(t)\right), \tag{4.50}$$

which requires that

$$\begin{vmatrix} \beta\sqrt{K}W^{EE} - 1 - \lambda\tau & -\beta\sqrt{K}W^{EI} \\ \beta\sqrt{K}W^{IE} & -\beta\sqrt{K}W^{II} - 1 - \lambda\tau \end{vmatrix} = 0 \tag{4.51}$$

and leads to

$$\begin{aligned} \lambda_{1,2} \quad &= \quad \frac{\beta\sqrt{K}\left(W^{EE} - W^{II}\right) - 2}{2\tau} \\ &\pm \quad \frac{1}{\tau}\sqrt{\left(\frac{\beta\sqrt{K}\left(W^{EE} - W^{II}\right) - 2}{2}\right)^2 - \beta^2 K\ W^{IE}W^{EI}} \\ \xrightarrow[K\to\infty]{} \quad &\frac{\beta\sqrt{K}}{\tau}\left[\frac{W^{EE} - W^{II}}{2} \pm \sqrt{\left(\frac{W^{EE} - W^{II}}{2}\right)^2 - \ W^{IE}W^{EI}}\right] \\ &= \quad \frac{\beta\sqrt{K}}{\tau}\left[\frac{W^{EE} - W^{II}}{2}\right]\left[1 \pm \sqrt{\left(1 - 4\frac{W^{IE}W^{EI}}{(W^{EE} - W^{II})^2}\right)}\right]. \end{aligned} \tag{4.52}$$

The system is stable only if the real part of $\lambda_{1,2} < 0$. This implies

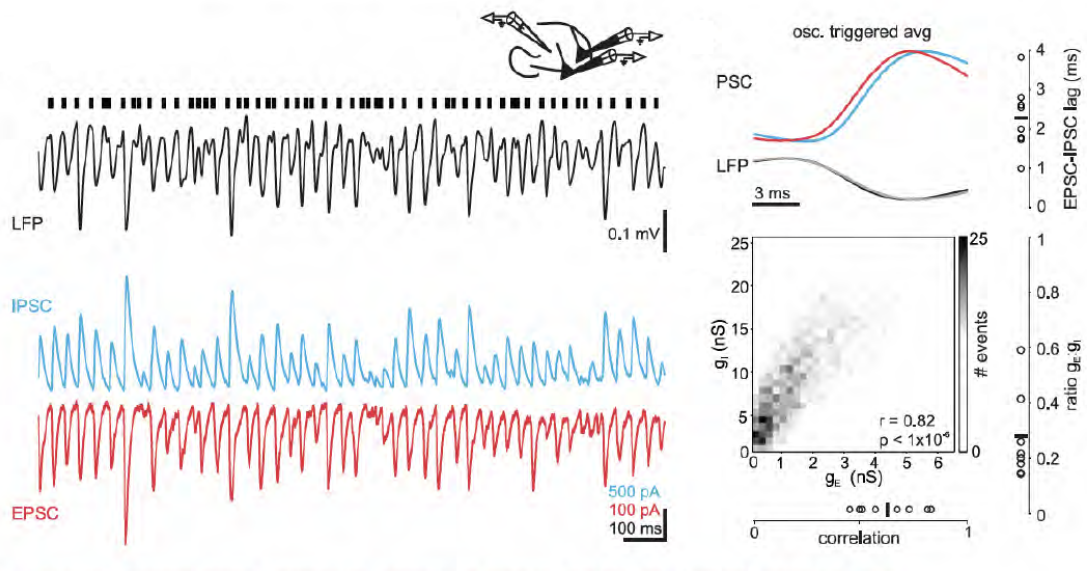
$$W^{II} > W^{EE}, \tag{4.53}$$

which is a prediction for connectomic analysis. We note that, by construction, $W^{IE}W^{EI} > 0$. The response time of the system is shortened by a factor of $\sqrt{K}$, i.e.,

$$\frac{\tau}{\beta} \to \frac{\tau}{\beta\sqrt{K}}O(1). \tag{4.54}$$

The change in recovery speed of the network has not been properly measured. But a sudden jump in the excitation of cortical input leads to an observed time-constant of about 10 ms (Figure 21). Unfortunately this is not very different from estimates for isolated neurons and thus the dynamics of the balanced still is a topic under analysis.

Figure 21: Relaxation of the signal in V1 cortical neurons after shut-down of thalamus. From Reinhold, Lien and Scanziani 2015



Box 1. Thermal noise starting with Johnson Noise formula

The expression for Johnson noise is

$$\delta V = \sqrt{\frac{4k_B T \Delta f}{G}}. \tag{4.55}$$

But the bandwidth for an RC circuit is

$$\begin{aligned} \Delta f &= \int_0^\infty df \, \frac{1}{1 + (2\pi f (C/G))^2} \\ &= \frac{G}{4C}. \end{aligned} \tag{4.56}$$

so

$$\delta V = \sqrt{\frac{k_B T}{C}}. \tag{4.57}$$

Box 2. Transfer function

Typically one measures two time series in recording from a neuron in an animal

- $V^k_{app}(t)$ is the applied stimulus or motor output for the k-th trial.
- $S^k_{meas}(t) = \sum_{spike\ times} \delta(t - t^k_r)$ is the measured spike time for the k-th trial.

How well can we reconstruct the stimulus from the spikes" We use the measured information to construct a linear filter that allows us to predict the stimulus for an unknown spike train.

In a sense. This idea is old, although it came into use only at the time of WW II, when there was a big push at the MIT Radar Laboratory to formulate the mathematics of optimal filtering and prediction. The procedure is as follows:

- We define $T(t)$ as the sought after transfer function.
- We define $V^k_{pred}(t)$ as the predicted stimulus for the k-th trial, based on the measured spike train, where

$$\begin{aligned} V^k_{pred}(t) &= \int_{-\infty}^{t} dt' T(t-t') S^k_{meas}(t') \\ &= \int_{-\infty}^{t} dt' T(t-t') \sum_s \delta(t' - t^k_S) \\ &= \sum_{spike\ times,\ S} \int_{-\infty}^{t} dt' T(t-t') \delta(t' - t^k_S) \\ &= \sum_S T(t - t^k_S) \end{aligned} \tag{4.58}$$

  is the predicted output, given by a convolution integral (Figure 22).

To determine $T(t)$, we minimize the difference between the actual and the predicted stimulus, averaged over all trials and time, i.e.,

$$\begin{aligned} Error &= \sum_k \int dt \left(V^k_{pred}(t) - V^k_{app}(t) \right)^2 \\ &= \sum_k \int dt \left(\int_{-\infty}^{t} dt' T(t-t') S^k_{meas}(t') - V^k_{app}(t) \right)^2 \end{aligned} \tag{4.59}$$

The error is computed in terms of measured quantities, except for $T(t)$, which we find by the criteria that we choose $T(t)$ to minimize the error. This is much easier to solve in the frequency domain, where convolutions turn into products. We consider the Fourier transformed variables:

$$V^k_{app}(t) \iff \tilde{V}^k_{app}(f) \tag{4.60}$$
$$S^k_{meas}(t) \iff \tilde{S}^k_{meas}(f) \tag{4.61}$$
$$V^k_{pred}(t) \iff \tilde{V}^k_{pred}(f) \tag{4.62}$$
$$T(t) \iff \tilde{T}(f) \tag{4.63}$$
$$\tag{4.64}$$

where

$$\tilde{T}(f) = \int_{-\infty}^{\infty} dt e^{i2\pi ft} T(t) \tag{4.65}$$
$$T(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} df e^{-i2\pi ft} \tilde{T}(f) \tag{4.66}$$

so that (ignoring causality for the moment) the convolution becomes

$$\int_{-\infty}^{\infty} dt' T(t-t') X(t') = \tilde{T}(f) \tilde{X}(f) \tag{4.67}$$

and we recall Parseval's theorem, effectively a conservation of energy, i.e.,

$$\int_{-\infty}^{\infty} dt|T(t)|^2 = \int_{-\infty}^{\infty} df|\tilde{T}(f)|^2 \tag{4.68}$$

where$|\tilde{T}(f)|^2 = \tilde{T}(f)\tilde{T}^*(f)$. We put the above together to write:

$$\begin{aligned}
Error &= \sum_k \int df \left|\tilde{V}^k_{pred}(f) - \tilde{V}^k_{app}(f)\right|^2 \qquad (4.69)\\
&= \int df \sum_k \left(\left|\tilde{V}^k_{pred}(f) - \tilde{V}^k_{app}(f)\right|^2\right)\\
&= \int df \sum_k \left(\left|\tilde{T}(f)\tilde{S}^k_{meas}(f) - \tilde{V}^k_{app}(f)\right|^2\right)\\
&= \int df \sum_k \left(\tilde{T}(f)\tilde{S}^k_{meas}(f) - \tilde{V}^k_{app}(f)\right)\left(\tilde{T}^*(f)\tilde{S}^{k*}_{meas}(f) - \tilde{V}^{k*}_{app}(f)\right)\\
&= \int df \sum_k \Big(\tilde{T}(f)\tilde{T}^*(f)|\tilde{S}^k_{meas}(f)|^2 - \tilde{T}(f)\tilde{S}^k_{meas}(f)\tilde{V}^{k*}_{app}(f)\\
&\quad -\tilde{V}^k_{app}(f)\tilde{T}^*(f)\tilde{S}^{k*}_{meas}(f) + |\tilde{V}^k_{app}(f)|^2\Big)\\
&= \int df\tilde{T}(f)\tilde{T}^*(f)\sum_k |\tilde{S}^k_{meas}(f)|^2 - \int df\tilde{T}(f)\sum_k \tilde{S}^k_{meas}\tilde{V}^{k*}_{app}(f)\\
&\quad - \int df\tilde{T}^*(f)\sum_k \tilde{V}^k_{app}\tilde{S}^{k*}_{meas}(f) + \int df\sum_k |\tilde{V}^k_{app}(f)|^2
\end{aligned}$$

The next step is to minimize the error with respect to the transfer function. We compute the function derivative

$$\frac{\partial(Error)}{\partial\tilde{T}^*(f)} = 0 \tag{4.70}$$

so that

$$\begin{aligned}
0 &= \int df\tilde{T}(f)\sum_k |\tilde{S}^k_{meas}(f)|^2 - \int df\sum_k \tilde{V}^k_{app}\tilde{S}^{k*}_{meas}(f) \qquad (4.71)\\
&= \int df\left(\tilde{T}(f)\sum_k |\tilde{S}^k_{meas}(f)|^2 - \sum_k \tilde{V}^k_{app}\tilde{S}^{k*}_{meas}(f)\right)
\end{aligned}$$

The expression for $T(f)$ must be valid at each frequency. Thus the frequency representation of the transfer function is

$$\tilde{T}(f) = \frac{\sum_k \tilde{V}^k_{app}(f)\tilde{S}^{k*}_{meas}(f)}{\sum_k |\tilde{S}^k_{meas}(f)|^2} \tag{4.72}$$

This is the central result. For the case of measured signal that is a spike train,

$$\tilde{T}(f) = \frac{\sum_k \tilde{V}^k_{app}(f)\sum_s e^{i2\pi f t^k_S}}{\sum_k \sum_{S,S'} e^{i2\pi f(t^k_S - t^k_{S'})}} \tag{4.73}$$

In the time domain, this is just

$$T(t) = \frac{1}{2\pi}\int df e^{-i2\pi ft}\frac{\sum_k \tilde{V}^k_{app}(f)\sum_s e^{i2\pi f t^k_S}}{\sum_k \sum_{s,s'} e^{i2\pi f(t^k_S - t^k_{S'})}} \tag{4.74}$$

Ugly! But this has a simple form when the spike arrival times may be taken to be a random, e.g., Poisson variable. This occurs if the spike rate is not too high, so that the refractory period plays no role. In this case the denumerator is just

$$\sum_k \sum_{S,S'} e^{i2\pi f(t_S^k - t_{S'}^k)} \approx N \tag{4.75}$$

where N is the total number of spikes across all trials, and the numerator is just

$$\begin{aligned} \frac{1}{2\pi}\int df e^{-i2\pi ft} \sum_k \tilde{V}_{app}^k(f) \sum_S e^{i2\pi f t_S^k} &= \sum_k \sum_S \frac{1}{2\pi}\int df e^{-i2\pi f(t-t_S^k)} \tilde{V}_{app}^k(f) \\ &= \sum_k \sum_S V_{app}^k(t - t_S^k) \end{aligned} \tag{4.76}$$

Thus $T(t)$ is just the spike triggered average of the stimulus waveform, i.e.,

$$T(t) \approx \frac{1}{N} \sum_k \sum_S V_{app}^k(t - t_S^k) \tag{4.77}$$

and finally we see that all that happens is that the transfer function reports the waveform of the stimulus that is most likely to cause the neuron to fire.

Figure 22: Pictorial guide to the convolution of $v(t)$ with $u(t)$.