

Randomly connected network

Peter Latham, January 19, 2023

1 Background

Our goal is to understand the dynamics of networks of recurrently connected excitatory and inhibitory neurons. Loosely, this means: given single neuron properties and connectivity, predict the qualitative behavior of a network *without doing any computer simulations*. This is hard, if not impossible, to do in general, but we'll consider cases where it can be done.

We have in mind networks of the form

$$\tau_m \frac{dV_i}{dt} = f(V_i) - \sum_j A_{ij} g_j(t) (V_i - \mathcal{E}_j) \quad (1a)$$

$$\tau_s \frac{dg_j}{dt} = -g_j(t) + \sum_{j,k} \delta(t - t_j^k). \quad (1b)$$

Here $f_i(V)$ represents the single neuron dynamics, \mathcal{E}_j is the reversal potential associated with neuron j , $\delta(\cdot)$ is the Dirac δ -function, and t_j^k is the time of the k^{th} spike on neuron j . Note that we could write down much more complicated equations – for instance, we could include axonal delays and dendritic processing, the dynamics of g_j could be made more complicated, g_j could depend on i as well as j , and, of course, the single neuron dynamics could have additional variables associated with channels, resulting in Hodgkin-Huxley type dynamics. However, the level of complexity in Eq. (1) will be sufficient to get the main ideas across.

Given these equations, we would like to determine things like: the equilibrium (or equilibria if there are more than one), stability around that equilibria, dynamics in the case of unstable equilibria, and we might even want to predict the distribution of firing rates. We can't (yet) do any of that in general, but we can do some of it for some connectivity structures. "Some" means two or three, depending on how you count, but that may be enough to mainly understand how networks work before learning. Here we'll focus on randomly connected networks, and in the next couple of sections we'll "solve" the above equations in various limits.

2 Firing rate equations

A very phenomenological approach to studying networks is to assume that neurons are completely described by their firing rates, and that the firing rate of any one neuron is a

function of the firing rates of all the other neurons in the network. Given this assumption (which is not exactly true, but is not terribly false either), a reasonable model is that the “synaptic drive” to a neuron is a linear sum of the firing rates of its pre-synaptic neurons, and the firing rate of a postsynaptic neurons is a nonlinear function of the synaptic drive. This produces a model of the form

$$\nu_i = \phi(h_i + h_{x,i}) \tag{2a}$$

$$h_i = \sum_j B_{ij}\nu_j \tag{2b}$$

where ν_i is the firing rate of neuron i , $h_{x,i}$ is the external input to neuron i , B_{ij} is the connection strength from neuron j to neuron i , and ϕ is the gain function. The gain function is typically approximately sigmoidal – zero when its argument is negative and large, and around 100 Hz when its argument is positive and large. We could have let ϕ depend on index, i , to make it different for every neuron, but that would have complicated notation without adding anything conceptual. Note that we’re using B_{ij} for the weights instead of A_{ij} (the latter being the weights that appeared in Eq. (1)). That’s because the weights that connect firing rates to firing rates are different from those that connect input spikes to voltage.

In the following we’ll solve Eq. (2) for two kinds of randomly connected networks: those that violate Dale’s law, and those that don’t. (Recall that Dale’s law tells us that a neuron makes either excitatory connections or inhibitory connections, but never both.) In Sec. 3, we’ll go back to Eq. (1), and take into account time dependence.

2.1 Networks that violate Dale’s law

We’ll start by considering a network of N neurons in which the weights, B_{ij} , are drawn *iid* from a single distribution. Although this is inconsistent with Dale’s law, it illustrates most of the techniques that we’ll use. To be somewhat realistic, we’ll consider sparse connectivity, and we’ll let the probability of a connection be K/N . This corresponds to an average of K connections per neurons, and it means that on average a fraction $1 - K/N$ of the B_{ij} are zero. Typically K/N is on the small side – about 1/10. To be honest, sparse connectivity adds very little conceptually, so on first reading it’s OK to mentally set K to N , which simplifies some of the analysis.

Given that connectivity is random, it turns out that all we need are the mean and variance of the elements of B_{ij} . For now we’ll leave those arbitrary. We will, however, introduce some scaling; we’ll let

$$B_{ij} = \frac{W_{ij}}{K^{1/2}}. \tag{3}$$

With this scaling, the synaptic drive, Eq. (2b), becomes

$$h_i = \frac{1}{K^{1/2}} \sum_{j=1}^N W_{ij} \nu_j. \quad (4)$$

(In what follow, all sums will be from 1 to N , so we'll suppress those limits below.) It's not hard to see why we introduced the factor of $1/K^{1/2}$: the sum on the right hand side of Eq. (4) consists of about K terms, and so different realizations of that sum (different values of i) will have a spread that scales as $K^{1/2}$; dividing by $1/K^{1/2}$ makes that spread independent of K . As we'll see below, the $\mathcal{O}(1)$ spread in synaptic drive is critical for producing a reasonable distribution of firing rates.

Our goal is to solve Eq. (2a) with the synaptic drive given in Eq. (4). Here “solve” doesn't mean find the firing rate of every neuron – because connectivity is random, the identity of any one neuron doesn't have much meaning. Instead, “solve” means “find the distribution over firing rates”.

To do that, we use essentially one approximation, which is so important that we'll highlight it in red:

Whenever we see a sum over a large number of indices, we treat it as a Gaussian random variable.

The right hand side of Eq. (4) qualifies as such a sum (it has about K nonzero terms, and in the brain K is about 1,000). Thus, we treat h_i as a Gaussian random variable. Note that it's Gaussian with respect to index i , meaning that if we were to make a histogram of all the h_i 's, it would look Gaussian.

The reason this is so important is that with the Gaussian assumption, all we need are the mean and variance of h to calculate the distribution over the firing rates, $P(\nu)$ (assuming we know the distribution over the input, h_x); that's given by the usual expression

$$P(\nu) = \int dh P(h) \int dh_x P(h_x) \delta(\nu - \phi(h + h_x)). \quad (5)$$

Here $P(h)$ is a Gaussian distribution (with mean and variance that we have to compute, which we'll do next), and $P(h_x)$ is the distribution over input, h_x (which we take to be independent of h). (If this doesn't make sense, you can also get the firing rate by sampling: sample h from a Gaussian distribution, sample h_x from it's distribution, set ν to $\phi(h + h_x)$, and repeat.) This doesn't, of course, tell us how to get the mean and variance of h , but that turns out to be a straightforward application of the central limit theorem, as we show now.

Our starting point is to break W_{ij} into a mean and fluctuating piece,

$$W_{ij} = \langle W \rangle + \delta W_{ij} \quad (6)$$

where $\langle W \rangle$ is the true mean of the weights (unless specified otherwise, angle brackets denote an average over the true distribution of whatever is inside them). We should think of this equation as defining δW_{ij} . Combining this relationship with Eq. (4), we see that

$$h_i = \frac{N}{K^{1/2}} \langle W \rangle \bar{\nu} + \frac{1}{K^{1/2}} \sum_j \delta W_{ij} \nu_j \quad (7)$$

where in general the k^{th} empirical moment of the firing rate is given by

$$\bar{\nu}^k \equiv \frac{1}{N} \sum_i \nu_i^k. \quad (8)$$

We can get $\bar{\nu}$ from this expression by setting $k = 1$.

The first term in Eq. (7) appears to be proportional to N/\sqrt{K} . However, that's not the actual scaling, since most of the elements of W_{ij} are zero. Defining W (with no subscripts) to be the true average of the nonzero terms in W_{ij} , we have (because of the sparse connectivity)

$$\langle W \rangle = \frac{K}{N} W. \quad (9)$$

Consequently, Eq. (7) becomes

$$h_i = K^{1/2} W \bar{\nu} + \frac{1}{K^{1/2}} \sum_j \delta W_{ij} \nu_j. \quad (10)$$

As promised above, we're going to treat the second term in this expression as a Gaussian random variable with respect to the index i . Its mean is zero, so we just need its variance, denoted σ_h^2 ,

$$\sigma_h^2 = \frac{1}{N} \sum_i \left(\frac{1}{K^{1/2}} \sum_j \delta W_{ij} \nu_j \right)^2 \quad (11)$$

Turning the square of the single sum into a double sum, and rearranging terms slightly, we have

$$\sigma_h^2 = \frac{1}{N} \sum_{jj'} \nu_j \nu_{j'} \left(\frac{1}{K} \sum_i \delta W_{ij} \delta W_{ij'} \right). \quad (12)$$

Separating this into terms with $j = j'$ and $j \neq j'$ yields

$$\sigma_h^2 = \frac{1}{N} \sum_j \nu_j^2 \left(\frac{1}{K} \sum_i \delta W_{ij}^2 \right) + \frac{1}{N} \sum_{j \neq j'} \nu_j \nu_{j'} \left(\frac{1}{K} \sum_i \delta W_{ij} \delta W_{ij'} \right). \quad (13)$$

The first term is $\mathcal{O}(1)$ (as we'll see shortly). The second term consists of a sum containing N^3 zero mean, uncorrelated random variables (OK, assumed-to-be-uncorrelated random variables, since the firing rates are actually correlated with the weights; why we can ignore those correlations is nontrivial), so it scales as $N^{3/2}\bar{\nu}^2\text{Var}[W]/NK$. It turns out that $\text{Var}[W]$ scales as K/N (for the same reason that the first term is $\mathcal{O}(1)$); consequently, the second term scales as $1/\sqrt{N}$. Thus, in the large N limit we can drop that term and focus on the first.

The sum over i in the first term can be expanded as

$$\frac{1}{K} \sum_i \delta W_{ij}^2 = \frac{1}{K} \sum_i \langle \delta W_{ij}^2 \rangle + \frac{1}{K} \sum_i (\delta W_{ij}^2 - \langle \delta W_{ij}^2 \rangle) \quad (14)$$

where, as above, the angle brackets denote an average over the true distribution of weights. The second term consists of N zero mean, uncorrelated random variables, so it scales as \sqrt{N}/K at worst, although it turns out to scale as $1/\sqrt{K}$ (which is not totally easy to show, but not so hard either). Given that scaling, in the large K limit the second term is smaller than the first term, so we can ignore it. To compute the first term, we just need $\langle \delta W_{ij}^2 \rangle$, which is given by

$$\langle \delta W_{ij}^2 \rangle = \langle W^2 \rangle - \langle W \rangle^2 = \frac{K}{N} (W^2 + \sigma_{w,\text{nonzero}}^2) - \frac{K^2}{N^2} W^2 \quad (15)$$

where $\sigma_{w,\text{nonzero}}^2$ is the variance of the nonzero weights. Rearranging terms slightly gives

$$\langle \delta W_{ij}^2 \rangle = \frac{K}{N} \left(\sigma_{w,\text{nonzero}}^2 + \left(1 - \frac{K}{N}\right) W^2 \right). \quad (16)$$

Consequently, to leading order in K ,

$$\frac{1}{K} \sum_i \delta W_{ij}^2 = \sigma_{w,\text{nonzero}}^2 + \left(1 - \frac{K}{N}\right) W^2. \quad (17)$$

Inserting this into Eq. (13) then yields

$$\sigma_h^2 = \sigma_w^2 \bar{\nu}^2 \quad (18)$$

where

$$\sigma_w^2 \equiv \sigma_{w,\text{nonzero}}^2 + \left(1 - \frac{K}{N}\right) W^2 \quad (19)$$

and, recall, $\bar{\nu}^2$ is given in Eq. (8). Again, on first reading, set $K = N$; that will reduce a lot of the algebra without any conceptual loss.

So far the analysis has been reasonably straightforward. But we'll now take a leap and treat the second term in Eq. (10) (whose variance we just computed) as a Gaussian random variable. We thus write

$$h_i = K^{1/2} W\bar{\nu} + \sigma_h \xi_i \quad (20)$$

where ξ_i is a zero mean, unit variance Gaussian random variable. Inserting this into Eq. (2a) gives us

$$\nu_i = \phi \left(K^{1/2} W\bar{\nu} + \sigma_h \xi_i + h_{x,i} \right) . \quad (21)$$

The right hand side depends on the first and second moments of the firing rate, which we don't know. However, we can find them self consistently by simply computing the empirical moments,

$$\overline{\nu^k} = \frac{1}{N} \sum_i \phi^k \left(K^{1/2} W\bar{\nu} + \sigma_h \xi_i + h_{x,i} \right) . \quad (22)$$

In the large N limit, samples can be replaced by averages over distributions, which gives us

$$\overline{\nu^k} = \int D\xi \int dh_x P(h_x) \phi^k \left(K^{1/2} W\bar{\nu} + \sigma_h \xi + h_x \right) \quad (23)$$

where, recall, $P(h_x)$ is the distribution over the input, h_x , and D (which can operate on any variable) is defined via

$$D\xi \equiv \frac{e^{-\xi^2/2}}{\sqrt{2\pi}} . \quad (24)$$

It's worth commenting here on the scaling with K , which we'll assume is large. First, the mean drive – the first term inside ϕ in Eq. (21) – is $\mathcal{O}(K^{1/2})$. Thus, to keep the gain function from saturating (so that the mean firing rate is neither 0 nor ν_{\max}), either W must be $\mathcal{O}(1/K^{1/2})$ or the external input, $h_{x,i}$, must balance, almost perfectly, the $K^{1/2}$ term (although in the latter case, the mean weight must be negative; see the discussion of stability below). Assuming that one of these conditions holds, so we manage to get the mean firing rate between 0 and ν_{\max} , then the distribution in firing rate is determined by σ_h . To have a reasonable distribution – neither too narrow nor too broad (in the latter case bimodal) – σ_h must be $\mathcal{O}(1)$. Only the $K^{-1/2}$ scaling of the weights ensures this.

Equation (23) gives us two equations ($k = 1, 2$) and two unknowns ($\bar{\nu}$ and $\overline{\nu^2}$). In general it must be solved self-consistently. However, to gain a qualitative understanding of the behavior of this equation, we don't really have to solve both. Instead, imagine solving the equation for the second moment, $\overline{\nu^2}$, in terms of the first, $\bar{\nu}$. In that case, the equation for the mean firing rate becomes

$$\bar{\nu} = \int D\xi \int dh_x P(h_x) \phi \left(K^{1/2} W\bar{\nu} + \sigma_h(\bar{\nu})\xi + h_x \right) . \quad (25)$$

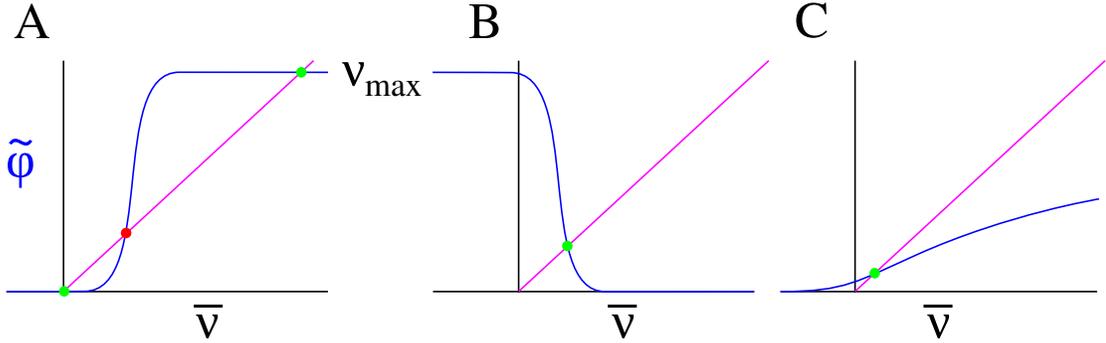


Figure 1. Plots of $\tilde{\phi}(K^{1/2}W\bar{v} + \langle h_x \rangle)$ (blue curve) versus \bar{v} for various values of K and W . The 45 degree line is shown in magenta and the red and green filled circles are the fixed points; green are stable and red are unstable. **A)** Positive mean weight, W , and relatively large K . **B)** Negative mean weight, W , and relatively large K . **C)** Positive mean weight, W , and relatively small K .

The integral over ξ is a convolution, so it simply smooths the gain function. Assuming the distribution of h_x is more or less Gaussian as well (probably all we need is that it has a single peak), the integral over h_x also simply smooths the gain function. We can, therefore write

$$\bar{v} = \tilde{\phi}(K^{1/2}W\bar{v} + \langle h_x \rangle) \quad (26)$$

where $\tilde{\phi}$, which is defined implicitly via Eq. (25), is a smoothed version of ϕ , with the degree of smoothing scaling with σ_h and the width of $P(h_x)$. Note that the degree of smoothing depends on \bar{v} , a dependence we suppress for clarity. Although we don't know exactly what $\tilde{\phi}$ looks like, for most (if not all) of our analysis its precise shape won't matter. This will become a lot more clear when we consider networks that obey Dale's law, which we do in the next section.

To find the solutions to Eq. (26), we can simply plot the right hand side versus \bar{v} and look for intersections with the 45 degree line. We do that in Fig. 1 for various values of the parameters. The blue curves in these plots are $\tilde{\phi}$, the magenta line is the 45 degree line, and the red and green filled circles are the fixed points. Panels A and B are plots with relatively large K for, respectively, positive and negative mean weight, W . Panel C shows a plot with positive mean weight but a much smaller value of K . Note that in all three plots the blue curve could be shifted in either direction; for panels A and B, a sufficiently large shift will eliminate two of the fixed points.

While this analysis can tell us about equilibria, it gives no hint about stability. For that we add, in a very hacky way, time dependence,

$$\tau \frac{d\bar{v}}{dt} = \tilde{\phi}(K^{1/2}W\bar{v} + \langle h_x \rangle) - \bar{v}. \quad (27)$$

This certainly has the right flavor: if, for instance, we suddenly increased synaptic drive, that would cause the firing rate to go up, but with a delay (here determined by the time constant, τ). However, although this equation has the right flavor, and gives reasonable intuition, I have no idea how to derive it from first principles.

To determine stability, we linearize the right hand side around a fixed point, denoted $\bar{\nu}^*$. Writing $\bar{\nu} = \bar{\nu}^* + \delta\bar{\nu}$, we have, in the limit of infinitesimal $\delta\bar{\nu}$,

$$\tau \frac{d\delta\bar{\nu}}{dt} = \left(K^{1/2} W \tilde{\phi}'(K^{1/2} W \bar{\nu}^* + \langle h_x \rangle) - 1 \right) \delta\bar{\nu} \quad (28)$$

where a prime denotes a derivative. The fixed point is stable if the term inside the parentheses is negative and unstable if it's positive. Thus, the condition for stability is that the slope of $\tilde{\phi}$ with respect to $\bar{\nu}$ is less than one. The easiest way to determine this is to simply plot $\tilde{\phi}(K^{1/2} W \bar{\nu} + \langle h_x \rangle)$ versus $\bar{\nu}$, and look at the slope at an equilibrium. If the slope is less than 1 the equilibrium is stable; if it's greater than 1 it's unstable. With this approach, it's easy to see that the red filled circle in Fig. 1 correspond to an unstable equilibrium and the green filled circles to stable equilibria. Note that in the large K limit, the intermediate intersection (if it exists) can be stable only if $W < 0$; that is, only if the neurons are inhibitory on average.

Finally, note that the intermediate intersection, which occurs at $\bar{\nu} \approx -\langle h_x \rangle / K^{1/2} W$, is $\mathcal{O}(1)$ only if $\langle h_x \rangle \propto K^{1/2}$. This is a general result, which we'll see again in the next section: for the external input to effect mean the firing rate, it needs to be $\mathcal{O}(K^{1/2})$.

2.2 Networks that obey Dale's law

The analysis in the previous section ignored a salient feature of the brain: Dale's law, which says (more or less) that any particular neuron makes either excitatory connections or inhibitory connections, but never both. We thus need to redo the analysis, but with extra indices: E for excitatory and I for inhibitory. The analysis is essentially the same, although it's more complicated because we have to keep track of which neurons are excitatory and which are inhibitory. However, the network is much more interesting.

We'll start by rewriting Eq. (2) as

$$\nu_{\alpha i} = \phi \left(\sum_{\beta} h_i^{\alpha} + \sqrt{K} h_{\alpha x} + \delta h_i^{\alpha x} \right) \quad (29a)$$

$$h_i^{\alpha} = \frac{1}{K^{1/2}} \sum_{\beta, j} W_{ij}^{\alpha\beta} \nu_{\beta j} \quad (29b)$$

where α and β can be either E or I and the \sqrt{K} scaling in front of $h_{\alpha x}$ was inspired by the discussion at the end of the previous section. Note that we're using the convention that the inhibitory weights are negative, which is different than what I used in class. But it makes the equations somewhat neater.

Again, all we need are the mean and variance of $h_{\alpha i}$. Using exactly the same analysis as above, we find that

$$\langle h^\alpha \rangle = K^{1/2} \sum_{\beta} W_{\alpha\beta} \bar{\nu}_\beta \quad (30a)$$

$$\text{Var}[h^\alpha] = \sum_{\beta} \sigma_{\alpha\beta}^2 \bar{\nu}_\beta^2 \quad (30b)$$

where the moments of the firing rates have definitions analogous to Eq. (8) and

$$W_{\alpha\beta} \equiv \text{mean of nonzero weights from neurons of type } \beta \text{ to neurons of type } \alpha \quad (31a)$$

$$\sigma_{\alpha\beta}^2 \equiv \sigma_{\alpha\beta, \text{nonzero}}^2 + \left(1 - \frac{K}{N}\right) W_{\alpha\beta}^2. \quad (31b)$$

As above, $\sigma_{\alpha\beta, \text{nonzero}}^2$ is the variance of the nonzero weights projecting from neurons of type β to neurons of type α . For simplicity, we assumed that the number of excitatory and inhibitory neurons, denoted N , is the same.

Inserting Eq. (30) into Eq. (29a) gives us

$$\nu_{\alpha i} = \phi \left(K^{1/2} \sum_{\beta} W_{\alpha\beta} \bar{\nu}_\beta + \left(\sum_{\beta} \sigma_{\alpha\beta}^2 \bar{\nu}_\beta^2 \right)^{1/2} \xi_i^\alpha + \delta h_i^{\alpha x} \right) \quad (32)$$

where ξ_i^α is a zero mean, unit variance Gaussian random variable. As for the non-Dale's law case, to have an $\mathcal{O}(1)$ spread in firing rates, the variances, $\sigma_{\alpha\beta}^2$, must be $\mathcal{O}(1)$. And again, this happens only when the weights scale as $K^{-1/2}$.

Averaging Eq. (32) over index, i , assuming, for simplicity, that $\delta h_i^{\alpha x}$ is a zero mean Gaussian random variable with variance $\sigma_{\alpha x}^2$, and retracing the steps in the previous section that led to Eq. (23), we arrive at the mean field equations

$$\bar{\nu}_\alpha^k = \int D\xi \phi^k \left(K^{1/2} \left(\sum_{\beta} W_{\alpha\beta} \bar{\nu}_\beta + h_{\alpha x} \right) + \left(\sum_{\beta} \sigma_{\alpha\beta}^2 \bar{\nu}_\beta^2 + \sigma_{\alpha x}^2 \right)^{1/2} \xi \right). \quad (33)$$

This corresponds to equations for the first ($k = 1$) and second ($k = 2$) moments of the excitatory and inhibitory firing rates. Because we have both excitatory and inhibitory neurons, we have four equations.

As above, we can think of solving for the second moments in terms of the first, then performing the integrals over ξ . Those integrals just smooth the gain functions, leading to

$$\bar{\nu}_\alpha = \tilde{\phi} \left(K^{1/2} \left(\sum_{\beta} W_{\alpha\beta} \bar{\nu}_\beta + h_{\alpha x} \right) \right). \quad (34)$$

As above, the degree of smoothing depends on the $\bar{\nu}_\alpha$, and as above, we suppress that dependence for clarity.

Note that in the large K limit things greatly simplify. That's because the term with $K^{1/2}$ in Eq. (33) dominates. So in that limit we can find the mean firing rates just by solving a set of linear equations,

$$\sum_{\beta} W_{\alpha\beta} \bar{\nu}_{\beta} + h_{\alpha x} = 0. \quad (35)$$

These equations are valid so long the solution is such that $\bar{\nu}_{\beta}$ lies between 0 and ν_{\max} for β equals both E and I (for simplicity we take ν_{\max} to be the same for excitatory and inhibitory neurons). If $\bar{\nu}_{\beta}$ is outside this range, it should be set to either 0 or ν_{\max} , whichever is appropriate. Taking this into account leads to the famous van Vreeswijk and Sompolinsky nullclines [1], and they do a pretty good job describing the mean firing rates. The effect of the so-called ‘‘quenched noise,’’ the ξ -related term in Eq. (33), is simply to induce a spread in firing rates. But from a conceptual point of view it doesn't add that much

As in the non-Dale's law case, the mean field equations don't tell us anything about either dynamics or stability. For that we have to add time dependence,

$$\tau_{\alpha} \frac{d\bar{\nu}_{\alpha}}{dt} = \tilde{\phi} \left(K^{1/2} \left(\sum_{\beta} W_{\alpha\beta} \bar{\nu}_{\beta} + h_{\alpha x} \right) \right) - \bar{\nu}_{\alpha}. \quad (36)$$

Again this is a total hack. But again it provides a pretty good picture of what's actually going on.

As we'll see below, although K is big (it's about 1,000), the weights are small. We thus often drop the explicit K -dependence, and write

$$\tau_E \frac{d\bar{\nu}_E}{dt} = \psi_E(\bar{\nu}_E, \bar{\nu}_I) - \bar{\nu}_E \quad (37a)$$

$$\tau_I \frac{d\bar{\nu}_I}{dt} = \psi_I(\bar{\nu}_E, \bar{\nu}_I) - \bar{\nu}_I \quad (37b)$$

where we have implicitly taken into account the smoothing associated with the quenched noise. These are a minor modification to the famous Wilson-Cowan equations [2] (see also [3], which contains a recipe for the construction of nullclines, along with lots of other tidbits); they're shown in what I believe is the relevant parameter regime in Fig. 2.

2.2.1 Summary for this section

Starting with a pretty much made-up model for the firing rates, we derived mean field equations for the first and second moments of the firing rate distribution. Once we solved those equations, we could find the distribution of firing rates. Our main approximation –

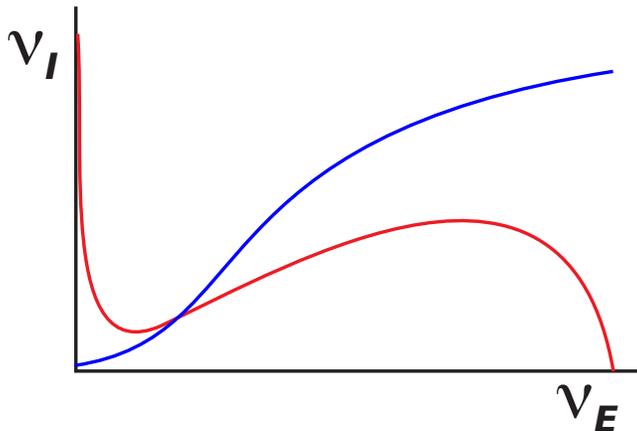


Figure 2. Wilson-Cowan nullclines (red for excitatory, blue for inhibitory) in the regime relevant for the brain.

one we'll use over and over – was to treat large sums as Gaussian random variables. This is clearly an approximation: for the sums to really be Gaussian, the elements have to be independent, which they aren't, since the firing rates depend on the weights. However, it turns out that for the networks we study, it's a pretty good approximation. The reasons are nontrivial, and we won't go into them here. We also made a second approximation, which was to turn sums into integrals. That, however, is almost always valid.

3 Time dependence

In the above analysis, we went directly to firing rate equations. We would very much like to derive those from Eq. (1), in a principled way. Unfortunately we can't, at least not very rigorously. However, we can gain insight into what's going on by treating the sum on the right hand side of Eq. (1a) as a time-dependent random variable. We'll start by singling out the synaptic drive in Eq. (1a),

$$h_i(t) \equiv - \sum_j A_{ij} g_j(t) (V_i - \mathcal{E}_j). \quad (38)$$

This is a little hard to deal with because V_i depends on t . We'll assume (without much thought – this is a major approximation) that V_i is independent of $g_j(t)$, and replace V_i with its average,

$$h_i(t) \approx - \sum_j A_{ij} g_j(t) (\langle V_i \rangle_t - \mathcal{E}_j) \quad (39)$$

where the subscript t on the angle brackets indicates a time average. Note that the weights are positive for excitatory neurons (for which $\mathcal{E} \approx 0$ mV whereas $\langle V_i \rangle_t \approx -60$ mV) and

negative for inhibitory neurons (for which $\mathcal{E} \approx -80$ mV). This means there are really two kinds of weights (excitatory and inhibitory) for which we need extra indices. Let us, therefore, define

$$W_{ij}^{\alpha\beta} = K^{1/2} A_{ij}(\mathcal{E}_j - \langle V_i \rangle_t), \quad i \in \alpha, j \in \beta \quad (40)$$

where the notation $i \in \alpha$ means neurons i is of type α (with, as above, α and β either E or I). We now need to put a superscript on h_i to tell us whether neuron i is excitatory or inhibitory, so we have

$$h_i^\alpha(t) \approx \frac{1}{K^{1/2}} \sum_{\beta,j} W_{ij}^{\alpha\beta} g_{\beta j}(t). \quad (41)$$

This is simply the time-dependent version of Eq. (29b), which possibly we could have written down immediately. With these approximations, Eq. (1a) becomes

$$\tau_m \frac{dV_{\alpha i}}{dt} = f(V_{\alpha i}) + \frac{1}{K^{1/2}} \sum_{\beta,j} W_{ij}^{\alpha\beta} g_{\beta j}(t). \quad (42)$$

This is a current-based, rather than the conductance-based, model.

As usual, we let

$$W_{ij}^{\alpha\beta} = \frac{K}{N} W_{\alpha\beta} + \delta W_{ij}^{\alpha\beta}. \quad (43)$$

As in Eq. (31a), $W_{\alpha\beta}$ represents the mean of the nonzero weights. We'll assume that $g_{\beta j}(t)$ obeys Eq. (1b). Consequently, the time average of $g_{\beta j}(t)$ is the firing rate, which we can easily see by integrating both sides over a long period and noting that $g_{\beta j}(t)$ is bounded. We can, therefore, write

$$g_{\beta j}(t) = \nu_{\beta j} + \delta g_{\beta j}(t). \quad (44)$$

Combining this with the above definition of $W_{ij}^{\alpha\beta}$, Eq. (41) becomes

$$h_i^\alpha(t) = K^{1/2} \sum_{\beta} W_{\alpha\beta} (\bar{\nu}_\beta + \delta \bar{g}_\beta(t)) + \frac{1}{K^{1/2}} \sum_{\beta,j} \delta W_{ij}^{\alpha\beta} (\nu_{\beta j} + \delta g_{\beta j}(t)) \quad (45)$$

where

$$\delta \bar{g}_\beta(t) \equiv \frac{1}{N} \sum_j \delta g_{\beta j}(t) \quad (46)$$

represents the population-averaged fluctuations in synaptic drive.

The time-independent terms in Eq. (45) we recognize from our analysis of the static case in the previous section, so we just have to deal with the time-dependent terms $\delta\bar{g}_\beta(t)$ and $\delta g_{\beta j}(t)$. Let's start with the former. Because the synaptic drive must be $\mathcal{O}(1)$, $\delta\bar{g}_\beta(t)$ must obey

$$\sum_{\beta} W_{\alpha\beta} \delta\bar{g}_\beta(t) \sim \frac{1}{K^{1/2}}. \quad (47)$$

This is true for $\alpha = E$ and I ; consequently, the average fluctuations are small for both the excitatory and inhibitory populations,

$$\bar{g}_\beta(t) \sim \frac{1}{K^{1/2}}. \quad (48)$$

As an (important) aside, this assumes that any time-varying input is $\mathcal{O}(1)$. If it's $\mathcal{O}(K^{1/2})$, all bets are off! However, if the input is $\mathcal{O}(1)$, this tells us that we can ignore $\bar{g}_\beta(t)$; that is, we can ignore time-varying fluctuations in the population averaged firing rates.

This has a second, and more important, implication: on average, correlations are weak. To see why, combine Eq. (48) with (46) to write

$$\langle \delta\bar{g}_\beta(t) \delta\bar{g}_{\beta'}(t + \tau) \rangle_t = \frac{1}{N^2} \sum_{jj'} \langle \delta g_{\beta j}(t) \delta g_{\beta' j'}(t + \tau) \rangle_t \sim \frac{1}{K} \quad (49)$$

where, as above, the subscript t indicates a time average. Consequently, when $\beta \neq \beta'$ or $j \neq j'$, on average

$$\langle \langle \delta g_{\beta j}(t) \delta g_{\beta' j'}(t + \tau) \rangle_t \rangle \sim \frac{1}{K}. \quad (50)$$

The reason we care about correlations is that they affect the last term in Eq. (45) – the term with the neuron-dependent fluctuations, $\delta g_{\beta j}$. Since this term is zero mean, what's relevant is its autocorrelation, which we'll denote $c_{\alpha i}(\tau)$,

$$c_{\alpha i}(\tau) \equiv \frac{1}{K} \sum_{\beta\beta', jj'} \delta W_{ij}^{\alpha\beta} \delta W_{ij'}^{\alpha\beta'} \langle \delta g_{\beta j}(t) \delta g_{\beta' j'}(t + \tau) \rangle_t \quad (51)$$

where again the subscript t indicates a time average. As usual, we single out the terms with $\beta = \beta'$ and $j = j'$, yielding

$$\begin{aligned} c_{\alpha i}(\tau) &= \frac{1}{K} \sum_{\beta, j} \left(\delta W_{ij}^{\alpha\beta} \right)^2 \langle \delta g_{\beta j}(t) \delta g_{\beta j}(t + \tau) \rangle_t \\ &+ \frac{1}{K} \sum_{\beta \neq \beta', j \neq j'} \delta W_{ij}^{\alpha\beta} \delta W_{ij'}^{\alpha\beta'} \langle \delta g_{\beta j}(t) \delta g_{\beta' j'}(t + \tau) \rangle_t. \end{aligned} \quad (52)$$

The second term is $\mathcal{O}(N/K^2)\text{Var}[\delta W_{ij}^{\alpha\beta}]$, with the second factor of $1/K$ coming from Eq. (50). The variance of $\delta W_{ij}^{\alpha\beta}$ is $\mathcal{O}(K/N)$ (see Eq. (16)), so the second term is $\mathcal{O}(1/K)$. As usual, because K is large, we can ignore it. This means

$$c_{\alpha i}(\tau) \approx \frac{1}{K} \sum_{\beta, j} \left(\delta W_{ij}^{\alpha\beta} \right)^2 \langle \delta g_{\beta j}(t) \delta g_{\beta j}(t + \tau) \rangle_t. \quad (53)$$

Assuming, as usual, that the weights are independent of the $\delta g_{\beta j}$, we can pull the weights out of the sum, yielding

$$c_{\alpha i}(\tau) \approx \sum_{\beta} \sigma_{\alpha\beta}^2 C_{\beta}(\tau) \quad (54)$$

where, as in Eq. (31b), $\sigma_{\alpha\beta}^2$ is the variance of the nonzero weights, and we have defined

$$C_{\beta}(\tau) = \frac{1}{N} \sum_j \langle \delta g_{\beta j}(t) \delta g_{\beta j}(t + \tau) \rangle_t. \quad (55)$$

This is a nice result, because it tells us that the autocorrelation function is independent of neuron index, at least in the large K limit.

Putting all this together, we have

$$h_i^{\alpha}(t) = K^{1/2} \sum_{\beta} W_{\alpha\beta} (\nu_{\beta} + \delta \bar{g}_{\beta}(t)) + \left(\sum_{\beta} \sigma_{\alpha\beta}^2 \bar{\nu}_{\beta}^2 \right)^{1/2} \xi_i^{\alpha} + \sum_{\beta} \sigma_{\alpha\beta} \zeta_{\beta i}(t) \quad (56)$$

where, as usual, ξ_i^{α} is a zero mean, unit variance Gaussian random variable, and $\zeta_{\beta i}(t)$ is a time-dependent random function with autocorrelation

$$\langle \zeta_{\beta i}(t) \zeta_{\beta' j}(t + \tau) \rangle_t = \delta_{\beta\beta'} \delta_{ij} C_{\beta}(\tau). \quad (57)$$

The membrane potential thus obeys the equation

$$\tau_m \frac{dV_{\alpha i}}{dt} = f(V_i) + h_i^{\alpha}(t) + h_{\alpha x, i}(t) \quad (58)$$

where we have added external input, $h_{\alpha x, i}(t)$. What we should do is solve for $C_{\alpha}(\tau)$ self consistently. This, however, is a whole function (unlike in the static case, where we only had to solve for the first and second moments of the firing rates). There has been some work in this area [4, 5, 6] (see also <http://www.gatsby.ucl.ac.uk/~pel/tn/notes/dmf.pdf>), but it's hard to do rigorously for spiking neurons, and it's not clear it's a good idea to try, especially considering all the other interesting problems in neuroscience.

Finally, a word about the external input, $h_{\alpha x, i}(t)$. If it is $\mathcal{O}(1)$ and has a component that's independent of i , that component will have an $\mathcal{O}(K^{-1/2})$ effect on the firing rate. In essence, the balance conditions pin the firing rates. However, if the external input depends on neuron, i , then firing rates will be modified – as they must. And if the external input has an $\mathcal{O}(K^{1/2})$ component, that component must be independent of i ; otherwise, some neurons will fire at very high rates and others will be silenced. That point (along with a few others) was made in [7].

4 How high is connectivity?

Our analysis relied heavily on the large K assumption. Which seems reasonable, since in cortex K is on the order of 1,000. However, determining whether or not this actually is reasonable is tricky, as we'll see.

To address this issue, it turns out we need two things: the size of the weights, and the size of the autocorrelation function, Eq. (55). We'll start with the weights, which we can get from PSP (post-synaptic potential) sizes, which are more or less known. Our approach is to ask how much the membrane potential changes in response to a single presynaptic spike. Using Eq. (42), and assuming infinitely fast PSPs (remember, we just want a ballpark estimate), the change in membrane potential due to a spike on neuron j of type β is

$$\Delta V_{\alpha i} = \frac{W_{ij}^{\alpha\beta}}{K^{1/2}\tau_m} \quad (59)$$

(to derive this, use $g_{\beta j} = \delta(t)$). The left hand side is, on average, the PSP size, so we have

$$W_{ij}^{\alpha\beta} \sim K^{1/2}\tau_m V_{PSP}^{\alpha\beta} \quad (60)$$

where $V_{PSP}^{\alpha\beta}$ is the typical PSP size for a neuron of type α receiving a spike from a neuron of type β .

To understand the implications of this, we insert it into Eq. (56), and get (very approximately)

$$h_i^\alpha(t) \sim K \sum_{\beta} V_{PSP}^{\alpha\beta} (\tau_m \nu_\beta + \tau_m \delta \bar{g}_\beta(t)) + K^{1/2} \left(\left(\sum_{\beta} (\delta V_{PSP}^{\alpha\beta})^2 \tau_m^2 \bar{\nu}_\beta^2 \right)^{1/2} \xi_i^\alpha + \sum_{\beta} \delta V_{PSP}^{\alpha\beta} \tau_m \zeta_{\beta i}(t) \right) \quad (61)$$

where $\delta V_{PSP}^{\alpha\beta}$ is the standard deviation of $V_{PSP}^{\alpha\beta}$.

The first observation is that we now have a factor of τ_m in all our expressions. Since τ_m is on the order of 10 ms, for firing rates measured in Hz this corresponds to a factor of 1/100. Consequently, the term proportional to K scales as $10 \times$ firing rate in Hz \times PSP size. This isn't so bad – it's 20 PSPs for a firing rate of only 2 Hz. However, it does mean the gain functions aren't especially steep functions of either the excitatory or inhibitory mean firing rates.

On the upside, the quenched noise term (the term proportional to ξ_i^α) also has a factor of τ_m in it as well. Thus, the term proportional to K is 30 times larger than the quenched noise term, independent of the firing rate.

But what about the time dependent term, $\zeta_{\beta i}(t)$? Its autocorrelation function, $C_{\beta}(t)$, is given in Eq. (55). To compute this we need to know the statistics of the spike trains, which we don't. But we can get an estimate of how big it is by assuming that the neurons are firing with Poisson statistics and are uncorrelated. In that case, as we show in Sec. 5 (see in particular Eq. (78)),

$$C_{\beta}(\tau) = \frac{\bar{\nu}_{\beta}}{\tau_{\text{eff}}} G(\tau) \quad (62)$$

where τ_{eff} is the effective synaptic time constant, defined in Eq. (75) (and taken, for simplicity, to be the same for all neurons), and $G(\tau)$ is a function that is 1 when $\tau = 0$ and falls rapidly to zero when $|\tau| > \tau_{\text{eff}}$ (see Eq. (77) for its definition, and Eq. (82) for a specific example).

It appears that $C_{\beta}(\tau)$ scales as $1/\tau_{\text{eff}}$. Because τ_{eff} is approximately equal to the synaptic time constant, it can be small, implying that $C_{\beta}(\tau)$ can be large. However, this is a bit of an illusion: because of the term $\tau_m d/dt$ in Eq. (1a), the drive to the membrane potential is low-pass filtered, so what really matters is the size of the fluctuations averaged over the membrane time constant, τ_m . In fact, we can let $\tau_{\text{eff}} \rightarrow 0$, at which point the synaptic drive turns into white noise, but that doesn't mean fluctuations in the membrane potential are infinitely large.

The white noise limit is actually kind of nice, because in that limit $\zeta_{\beta i}(t) \rightarrow \bar{\nu}_{\beta}^{1/2} \eta_i(t)$ where $\eta_i(t)$ is white noise,

$$\langle \eta(t)\eta(t+\tau) \rangle_t = \delta(\tau). \quad (63)$$

In that limit (which is a worst case scenario, in the sense that nonzero synaptic time constant will decrease the temporal fluctuations),

$$\tau_m \zeta_{\beta i}(t) \rightarrow (\tau_m \bar{\nu}_{\beta})^{1/2} \tau_m^{1/2} \eta_i(t). \quad (64)$$

To determine how these temporal fluctuations affect the membrane potential, we'll isolate the temporal fluctuations by considering the equation

$$\tau_m \frac{dV}{dt} = -V + \tau_m^{1/2} \eta(t) \quad (65)$$

where $\eta(t)$ is white noise. This has the solution

$$V(t) = \int_{-\infty}^t \frac{dt'}{\tau_m^{1/2}} e^{-(t-t')/\tau_m} \eta(t'). \quad (66)$$

The resulting autocorrelation function is

$$C_V(\tau) \equiv \langle V(t)V(t+\tau) \rangle_t = \frac{1}{\tau_m} \int_{-\infty}^t dt' \int_{-\infty}^{t+\tau} dt'' e^{-(t-t')/\tau_m - (t+\tau-t'')/\tau_m} \langle \eta(t')\eta(t'') \rangle. \quad (67)$$

Assuming $\tau > 0$ (we'll find the autocorrelation function for negative τ by symmetry), we can use Eq. (63) to turn the average over the white noise into delta functions. After that the integrals are easy, and we arrive at

$$C_V(\tau) = \frac{e^{-|\tau|/\tau_m}}{2}. \quad (68)$$

Consequently, the term $\tau_m^{1/2}\eta_i(t)$ is effectively $\mathcal{O}(1)$, and so the ratio of temporal fluctuations to mean drive (compare the first and last terms in Eq. (61)) is

$$\frac{\text{temporal fluctuations}}{\text{mean drive}} \sim \frac{1}{(K\tau_m\bar{\nu}_\beta)^{1/2}}. \quad (69)$$

Because $K\tau_m \sim 10$ s, this isn't so small. Ken Miller has made a big deal of this [8], but the implications are mainly pretty minor; it just means the nullclines are not straight, as in van Vreeswijk and Sompolinsky's early work [1]; instead, they're curved, as in Fig. 2. It's still the case that K is large enough to use the central limit theory for the quenched noise, so that part is OK. The main implications as far as our derivation goes is that correlations may be larger than we think. Remember that we used large K to argue that correlations are $\mathcal{O}(1/K^{1/2})$ (see Eq. (50)). However, once we take into account the scaling factor of τ_m , that equation really should have been

$$\langle \delta g_{\beta j}(t) \delta g_{\beta' j'}(t + \tau) \rangle_t \sim \frac{1}{(K\tau_m\bar{\nu}_\beta)^{1/2}}, \quad (70)$$

which is not so small. And, in fact, correlations in the cortex tend to be on the order of 10%, and sometimes larger. However, to understand the average behavior of the networks, all of our analysis holds.

5 Temporal correlations for Poisson firing

Here we'll compute the autocorrelation function given in Eq. (55) when the neurons are uncorrelated and firing with Poisson statistics. There must be an easy way to do this, but I don't know what it is, so I'll use an insanely complicated method.

We'll start by writing (dropping the subscript β to reduce clutter)

$$\begin{aligned} C(\tau) &= \frac{1}{N} \sum_j \left\langle \sum_{l,m} \delta g(t - t_j^l) \delta g(\tau + t - t_j^m) \right\rangle \\ &= \frac{1}{N} \sum_j \left(\left\langle \sum_{l,m} g(t - t_j^l) g(\tau + t - t_j^m) \right\rangle - \nu_{\alpha j}^2 \right) \end{aligned} \quad (71)$$

where t_j^k is the time of the k^{th} spike on neuron j , the average is now over both t and the statistics of the spike trains, and we are assuming that the shape of the conductance change depends on α but not j . We'll perform this average in an interval of size T , which we'll eventually take to ∞ . Using the fact that the spike counts obey Poisson statistics, we may write

$$\left\langle \sum_{l,m} g(t - t_j^l) g(\tau + t - t_j^m) \right\rangle = \sum_{k=0}^{\infty} \frac{(\nu_{\alpha j} T)^k e^{-\nu_{\alpha j} T}}{k!} \int \prod_{n=1}^k \frac{dt_n}{T} \sum_{l,m} g(t - t_l) g(\tau + t - t_m) \quad (72)$$

where the integral is a k^{th} order temporal integral over the t_n , with each integral running from 0 to T . There are $k(k-1)$ terms that have $l \neq m$ and k terms that have $l = m$. In each case the time integrals are straightforward, and we arrive at

$$\left\langle \sum_{l,m} g(t - t_j^l) g(\tau + t - t_j^m) \right\rangle = \sum_{k=0}^{\infty} \frac{(\nu_{\alpha j} T)^k e^{-\nu_{\alpha j} T}}{k!} \left[\frac{k(k-1)}{T^2} + \frac{k}{T} \int dt g(t) g(t + \tau) \right] \quad (73)$$

where we have used the fact that $g(t)$ integrates to 1. Performing the sums over k , we arrive at

$$\left\langle \sum_{l,m} g(t - t_j^l) g(\tau + t - t_j^m) \right\rangle = \nu_j^2 + \nu_j \int dt g(t) g(t + \tau). \quad (74)$$

Defining the effective synaptic time constant as

$$\frac{1}{\tau_{\text{eff}}} \equiv \int dt g(t)^2 \quad (75)$$

(we show below that this is a sensible definition; see analysis starting with Eq. (80)), we arrive at

$$\left\langle \sum_{l,m} g(t - t_j^l) g(t + \tau - t_j^m) \right\rangle = \nu_j^2 + \frac{\nu_j}{\tau_{\text{eff}}} G(\tau) \quad (76)$$

where

$$G(\tau) \equiv \frac{\int dt g(t) g(t + \tau)}{\int dt g(t)^2}. \quad (77)$$

Finally, inserting this into Eq. (71), we arrive at the very simple expression

$$C(\tau) = \frac{1}{N} \sum_j \frac{\nu_j}{\tau_{\text{eff}}} = \frac{\bar{\nu}}{\tau_{\text{eff}}} G(\tau). \quad (78)$$

Note that $\int d\tau C(\tau)$ has a simple form: combining the above expression with Eq. (75), we see that

$$\int_{-\infty}^{\infty} d\tau C(\tau) = \bar{v} \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} dt g(t)g(t + \tau) = \bar{v}, \quad (79)$$

which follows because $g(t)$ integrates to 1. This is relevant for justifying Eq. (63).

To make sure this all makes sense, let's compute $G(\tau)$ for a decaying exponential, $g(t) = \Theta(t)e^{-t/\tau_s}/\tau_s$. In that case, assuming $\tau > 0$,

$$\int dt g(t)g(t + \tau) = \int_0^{\infty} \frac{dt}{\tau_s^2} e^{-t/\tau_s} e^{-(t+\tau)/\tau_s} = \frac{e^{-\tau/\tau_s}}{2\tau_s}. \quad (80)$$

Consequently, via Eq. (75),

$$\tau_{\text{eff}} = 2\tau_s, \quad (81)$$

and

$$G(\tau) = e^{-2|\tau|/\tau_{\text{eff}}} \quad (82)$$

(the absolute value sign comes from symmetry around $\tau = 0$).

References

1. C. van Vreeswijk and H. Sompolinsky, "Chaotic balanced state in a model of cortical circuits," *Neural Comput.*, vol. 10, no. 6, pp. 1321–1371, 1998.
2. H. Wilson and J. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophys. J.*, vol. 12, pp. 1–24, 1972.
3. P. Latham, B. Richmond, P. Nelson, and S. Nirenberg, "Intrinsic dynamics in neuronal networks. I. theory," *J. Neurophysiol.*, vol. 83, pp. 808–827, 2000.
4. A. Lerchner, C. Ursta, J. Hertz, M. Ahmadi, P. Ruffiot, and S. Enemark, "Response variability in balanced cortical networks," *Neural Comput.*, vol. 18, no. 3, pp. 634–659, 2006.
5. A. Grabska-Barwinska and P. E. Latham, "How well do mean field theories of spiking quadratic-integrate-and-fire networks work in realistic parameter regimes?," *J. Comp. Neurosci.*, vol. 36, pp. 469–481, 2014.
6. M. F. and O. S., "Intrinsically-generated fluctuating activity in excitatory-inhibitory networks," *PLoS Comp. Bio.*, vol. 13, no. 4, p. e1005498, 2017.

7. R. Rosenbaum, M. A. Smith, A. Kohn, J. E. Rubin, and B. Doiron, “The spatial structure of correlated neuronal variability,” *Nature Neurosci.*, vol. 20, pp. 107–114, 2017.
8. Y. Ahmadian and K. D. Miller, “What is the dynamical regime of cerebral cortex?”