

## Lecture notes on network dynamics

### 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*.

We have in mind network dynamics of the form

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

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

Here  $f_i(V, t)$  represents the single neuron dynamics,  $\mathcal{E}_j$  is the reversal potential associated with neuron  $j$ ,  $\delta(\cdot)$  is the Dirac  $\delta$ -function, and  $t_j^k$  is the time of the  $k^{\text{th}}$  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 this 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. Here we'll focus primarily on randomly connected networks, and in the next couple of sections we'll "solve" the above equations in various limits, and for various types of randomness. At the end, we'll consider more structured connectivity: that associated with Hopfield networks.

### 2 Firing rate equations

A very phenomenological approach to studying network dynamics is to assume that neurons are described totally 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 a network. Given this assumption (which we'll take a closer look at below), 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

$$\nu_i = \phi(h_i + I_i) \tag{2a}$$

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

where  $\nu_i$  is the firing rate of neuron  $i$ ,  $I_i$  is the input to neuron  $i$ ,  $B_{ij}$  is the connection strength from neuron  $j$  to neuron  $i$ , and  $\phi$  is the gain function. The later is typically sigmoidal – zero when its argument is negative and large, and around 100 Hz when its argument is positive and large. We could have let  $\phi$  depend on index,  $i$ , but that would have complicated notation without adding anything conceptual.

In the following we'll solve this equation 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.)

## 2.1 Networks that violate Dale's law

We'll start by considering a network in which the elements,  $B_{ij}$  are drawn *iid* from some distribution. Although this is inconsistent with biology, it illustrates most of the techniques that we'll use. To be 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 set  $K$  to  $N$ , which simplifies many of the expression.

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

$$h_i = \frac{1}{K^{1/2}} \sum_j W_{ij}\nu_j. \tag{4}$$

It's not hard to see why we use this scaling: the sum on the right hand side of Eq. (4) consists of about  $K$  terms, all of which scale as  $1/K^{1/2}$ ; it's spread is, therefore, about the size of the weights,  $W_{ij}$ , independent of the size of the network.

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. That’s because connectivity is random, so 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 synaptic drive,  $h_i$ , that appears in Eq. (4) qualifies as such a sum (it has about  $K$  nonzero terms). Thus, we treat  $h_i$  as a Gaussian random variable. Importantly, 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 nice thing about the Gaussian assumption is that all we need are the mean and variance of  $h$ . Once we have those, we can immediately find the distribution over the firing rates (assuming we know the distribution over the input,  $I_i$ ),

$$P(\nu) = \int dI P(I) \int dh p(h) \delta(\nu - \phi(h + I)). \quad (5) \quad \{\text{p\_nu}\}$$

Here  $P(h)$  is a Gaussian distribution, and  $P(I)$  is assumed known. (If this doesn’t make sense, you can also get the firing rate by sampling: sample  $h$  from a Gaussian distribution, sample  $I$  from its distribution, set  $\nu$  to  $\phi(h + I)$ , and repeat.)

So now we need the mean and variance of the Gaussian distribution over  $h$ . The mean is given by

$$\langle h \rangle = \frac{1}{N} \sum_i \frac{1}{K^{1/2}} \sum_j W_{ij} \nu_j = \frac{1}{K^{1/2}} \sum_j \nu_j \frac{1}{N} \sum_i W_{ij} = \frac{N}{K^{1/2}} \overline{W} \bar{\nu} \quad (6) \quad \{\text{hbar}\}$$

where  $\bar{\nu}$  is the population average firing rate ( $k = 1$  in Eq. (12)) and  $\overline{W}$  is the average connection strength. Strictly speaking,

$$\overline{W} \equiv \frac{1}{N^2} \sum_{ij} W_{ij}, \quad (7)$$

so in fact the right hand side of Eq. (6) is an approximation to  $\langle h \rangle$ . However, that approximation should be good in the large  $K$  limit, as fluctuations are  $\mathcal{O}(1/K^{1/2})$ . Note that for  $\langle h \rangle$  to be  $\mathcal{O}(1)$ ,  $\overline{W}$  must be  $\mathcal{O}(K^{1/2}/N)$ . This means that the nonzero weights (of which there are a fraction  $K/N$ ) are  $\mathcal{O}(1/K^{1/2})$ , which in turn means that the strengths of the nonzero connections are  $\mathcal{O}(1/K)$ .

The second moment of  $h$  is given by

$$\langle h^2 \rangle = \frac{1}{N} \sum_i \frac{1}{K} \sum_{jj'} W_{ij} \nu_j W_{ij'} \nu_{j'} = \frac{1}{K} \sum_{jj'} \nu_j \nu_{j'} \frac{1}{N} \sum_i W_{ij} W_{ij'}. \quad (8)$$

If  $j \neq j'$ , the sum over  $i$  is just  $\overline{W}^2$ ; otherwise, it's  $\langle W^2 \rangle$ , the second moment of the weights (in both we ignored  $O(1/K^{1/2})$  corrections). We thus have

$$\langle h^2 \rangle = \frac{1}{K} \sum_{jj'} \nu_j \nu_{j'} [\overline{W}^2 (1 - \delta_{jj'}) + \langle W^2 \rangle \delta_{jj'}] = \frac{N^2}{K} \overline{W}^2 \overline{\nu}^2 + \frac{N}{K} \text{Var}[W] \overline{\nu}^2 \quad (9)$$

where  $\overline{\nu}^2$  is the second moment of the firing rates ( $k = 2$  in Eq. (12)) and

$$\langle W^2 \rangle \equiv \frac{1}{N^2} \sum_{ij} W_{ij}^2. \quad (10)$$

Using Eq. (6) for the mean, we see that

$$\text{Var}[h] = \frac{N}{K} \text{Var}[W] \langle \nu^2 \rangle. \quad (11)$$

We now have expressions for the mean and variance of  $h$ , but they depend on the first and second moments of the firing rates, which we don't know. However, we can express those moments in terms of the mean and variance of  $h$ , giving us self-consistent equations. The equations will be nonlinear, and we will have to solve them numerically, but at least we don't have to do any network simulations.

The  $k^{\text{th}}$  moment of the firing rate is given by

$$\overline{\nu}^k \equiv \frac{1}{N} \sum_i \nu_i^k. \quad (12) \quad \{\text{moments}\}$$

To express this in terms of the distribution over  $h$ , we use Eq. (2a) to write

$$\overline{\nu}^k = \frac{1}{N} \sum_i \phi^k(h_i + I_i). \quad (13) \quad \{\text{moments}\}$$

In the large  $n$  limit, we can turn the sum on the right hand side into an integral over the distributions of  $h$  and  $I$ , giving us

$$\overline{\nu}^k = \int dI P(I) \int dh P(h) \phi^k(h + I). \quad (14)$$

We are assuming (as above) that  $h$  and  $I$  are independent.

It is convenient to make the definitions

$$\mu \equiv \frac{N}{K^{1/2}} \overline{W} \quad (15a)$$

$$\sigma^2 \equiv \frac{N}{K} \text{Var}[W], \quad (15b)$$

giving us

$$\overline{\nu^k} = \int dI P(I) \int Dx \phi^k \left( \mu \overline{\nu} + \sigma \overline{\nu}^{1/2} x + I \right) \quad (16) \quad \{\text{mf\_node}\}$$

where

$$Dx \equiv \frac{e^{-x^2/2}}{(2\pi)^{1/2}}. \quad (17)$$

Equation (16) actually corresponds to two equations: one for  $\overline{\nu}$  and one for  $\overline{\nu^2}$ ; these must be solved self-consistently. Once they are known, the full distribution of firing rates can be computed from Eq. (5).

## 2.2 Networks that obey Dale's law

The analysis in the previous ignored a salient feature of neuron in the brain: they almost always obey Dale's law, meaning 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 just because we have to keep track of which neurons are excitatory and which are inhibitory. We do, though, gain some new insight.

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

$$\nu_{\alpha i} = \phi(h_{\alpha i} + I_{\alpha i}) \quad (18a) \quad \{\text{ei}\}$$

$$h_{\alpha i} = \frac{1}{K^{1/2}} \sum_{\beta, j} W_{ij}^{\alpha\beta} \nu_{\beta j} \quad (18b) \quad \{\text{ei.b}\}$$

where  $\alpha$  and  $\beta$  can be either  $E$  or  $I$ . Again, all we need are the mean and variance of  $h_{\alpha i}$ . Using exactly the same analysis as above, we find that

$$\langle h_{\alpha} \rangle = K^{1/2} \sum_{\beta} W_{\alpha\beta} \overline{\nu}_{\beta} \quad (19a)$$

$$\text{Var}[h_{\alpha}] = \sum_{\beta} \sigma_{\alpha\beta}^2 \overline{\nu}_{\beta}^2 \quad (19b)$$

where

$$W_{\alpha\beta} \equiv \frac{1}{NK} \sum_{ij} W_{ij}^{\alpha\beta} \quad (20a)$$

$$\sigma_{\alpha\beta}^2 \equiv \frac{N}{K} \text{Var}[W_{ij}^{\alpha\beta}]. \quad (20b)$$

The moments of the firing rates have definitions analogous to Eq. (13).

Inserting this into Eq. (18), we arrive at the mean field equations

$$\bar{\nu}_\alpha^k = \int dI P(I_\alpha) \int Dx \phi^k \left( K^{1/2} \sum_\beta W_{\alpha\beta} \bar{\nu}_\beta + \left( \sum_\beta \sigma_{\alpha\beta}^2 \bar{\nu}_\beta^2 \right)^{1/2} x + I_\alpha \right). \quad (21) \quad \{\text{mf\_dale}\}$$

This corresponds to four equations that have to be solved self-consistently for the first ( $k = 1$ ) and second ( $k = 2$ ) moments of the excitatory and inhibitory firing rates.

Note that in the large  $K$  limit things greatly simplify. That’s because the term with  $K^{1/2}$  in Eq. (21) dominates. Assuming that the mean value of  $I_\alpha$  is  $\mathcal{O}(K^{1/2})$ , we can solve for the mean firing rates without doing any integrals,

$$\sum_\beta W_{\alpha\beta} \bar{\nu}_\beta + K^{-1/2} \bar{I}_\alpha = 0. \quad (22)$$

These lead to the famous van Vreeswijk and Sompolinsky nullclines, and they do a pretty good job describing the mean firing rates. The effect of the so-called “quenched noise,” the  $x$ -related term in Eq. (21), is twofold: it induces a spread in firing rates, and it smooths out the gain functions. But from a conceptual point of view it doesn’t add all that much

### 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 their first and second moments, which in turn could be used to find the distribution over the synaptic drive,  $h$ . Our main approximation – 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 – firing rates are correlated. 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 A (somewhat) more rigorous approach

While the above analysis illustrated the main techniques we use, it was relatively ad hoc. Here we do a better job deriving the firing rate equations, starting from Eq. (1). However, although we’ll be more rigorous, there will still be some loose ends. **To be continued ...**