

# What is the dynamical regime of cerebral cortex?

Yashar Ahmadian<sup>1,2</sup> and Kenneth D. Miller<sup>3</sup>

<sup>1</sup>Institute of Neuroscience, Departments of Biology and Mathematics, University of Oregon, OR, USA

<sup>2</sup>As of March 2020: Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge, UK

<sup>3</sup>Center for Theoretical Neuroscience, Swartz Program in Theoretical Neuroscience, Kavli Institute for Brain Science, and Dept. of Neuroscience, College of Physicians and Surgeons and Morton B. Zuckerman Mind Brain Behavior Institute, Columbia University, NY, USA

August 29, 2019

## Summary

Many studies have shown that the excitation and inhibition received by cortical neurons remain roughly balanced across many conditions. A key question for understanding the dynamical regime of cortex is the nature of this balancing. Theorists have shown that network dynamics can yield systematic cancellation of most of a neuron's excitatory input by inhibition. We review a wide range of evidence pointing to this cancellation occurring in a regime in which the balance is loose, meaning that the net input remaining after cancellation of excitation and inhibition is comparable in size to the factors that cancel, rather than tight, meaning that the net input is very small relative to the cancelling factors. This choice of regime has important implications for cortical functional responses, as we describe: loose balance, but not tight balance, can yield many nonlinear population behaviors seen in sensory cortical neurons, allow the presence of correlated variability, and yield decrease of that variability with increasing external stimulus drive as observed across multiple cortical areas.

In what regime does cerebral cortex operate? This is a fundamental question for understanding cerebral cortical function. The concept of a "regime" can be defined in various ways. Here we will focus on a definition in terms of the balance of excitation and inhibition: how strong are the excitation and inhibition that cortical cells receive, and how tight is the balance between them? As we will see, the answers to these questions have important implications for the dynamical function of cortex.

We first consider several more fundamental distinctions in cortical regime. First, neurons may fire in a regular or irregular fashion, where regular firing refers to emitting spikes in a more clock-like manner, while irregular firing refers to emitting spikes in a more random manner, like a Poisson process. Cortex appears to be in an irregular regime (Shadlen and Newsome, 1998; Softky and

Koch, 1993), though some areas are less irregular than others (Maimon and Assad, 2009). Second, neurons may fire in a synchronous regime, meaning with strong correlations between the firing of different neurons, or an asynchronous regime, meaning with weak (or no) correlations. Cortical firing, particularly in the awake state, generally appears to be in an asynchronous regime (Cohen and Kohn, 2011; Doiron et al., 2016; Ecker et al., 2014, 2010), although some conditions may show more synchronous firing (DeWeese and Zador, 2006; Poulet and Petersen, 2008; Stevens and Zador, 1998; Tan et al., 2014). Thus, we will take cortex to be in an asynchronous irregular regime. Brunel (2000) first defined conditions on networks of excitatory and inhibitory neurons that led them to operate in the asynchronous irregular regime.

A third distinction is whether a network goes to a stable fixed rate of firing for a fixed input (with noisy fluctuations about that fixed rate given noisy inputs), or whether it shows more complex behaviors, such as movement between multiple fixed points, oscillations, or chaotic wandering of firing rates. We will focus on the case of a single fixed point, which seems likely to reasonably approximate at least awake sensory cortex (see discussion in [Miller, 2016](#)). The fixed point is taken to be *stable*, meaning that the network dynamics cause firing rates to return to the fixed point levels after small perturbations. Finally, for a given fixed point, recurrent excitation may be strong enough to destabilize the fixed point in the absence of feedback inhibition; that is, if inhibitory firing were held frozen at its fixed point level, a small perturbation of excitatory firing rates would cause them to either grow very large or collapse to zero. In that case, the fixed point is stabilized by feedback inhibition, and the network is known as an inhibition-stabilized network (ISN). Alternatively, the recurrent excitation may be weak enough to remain stable even without feedback inhibition. A number of studies have found strong evidence that at least primary visual and auditory cortices are ISNs both at spontaneous ([Sanzeni et al., 2019](#)) and stimulus-driven ([Adesnik, 2017](#); [Kato et al., 2017](#); [Ozeki et al., 2009](#)) levels of activity.

Note that for some of the distinctions we describe between regimes, there is a sharp transition from one regime to the other, while for others the transition is gradual. We use the word “regime” in either case to describe qualitatively different network behaviors.

The assumption that cortex is in an irregularly-firing regime (as well as its operation as an ISN) strongly points to the need for some kind of balance between excitation and inhibition. Stochasticity of cellular and synaptic mechanisms ([Mainen and Sejnowski, 1995](#); [O’Donnell and van Rossum, 2014](#); [Schneidman et al., 1998](#)) and input correlations ([DeWeese and Zador, 2006](#); [Stevens and Zador, 1998](#)) may contribute to irregular firing. However, a number of authors have argued that, assuming inputs are un- or weakly-correlated, then irregular firing will arise if the mean input to cortical cells is sub- or peri-threshold, so that firing is induced by fluctuations from the mean rather than by the mean itself ([Amit and Brunel, 1997](#); [Troyer and Miller, 1997](#); [Tsodyks and Sejnowski, 1995](#); [van Vreeswijk and Sompolinsky, 1996](#)). This is referred to as the fluctuation-driven regime, as opposed to the mean-driven regime in which the mean

input is strongly suprathreshold and spiking is largely driven by integration of this mean input. The fluctuation-driven regime yields random, Poisson-process-like firing, because fluctuations are equally likely to occur at any time, whereas the mean-driven regime yields regular firing.

Given the strength of inputs to cortex (to be discussed below), the mean excitation received by a strongly-responding cell is likely to be sufficient to drive the cell near or above threshold. Therefore, for the mean input to be subthreshold, the mean inhibition is likely to cancel a significant portion of the mean excitation; that is, the excitation and inhibition received by a cortical cell should be at least roughly balanced ([Tsodyks and Sejnowski, 1995](#); [van Vreeswijk and Sompolinsky, 1996](#)). Consistent with the idea that inhibition balances excitation, many experimental investigations have suggested that cortical or hippocampal excitation and inhibition remain balanced or inhibition-dominated across varying activity levels ([Anderson et al., 2000](#); [Atallah and Scanziani, 2009](#); [Barral and Reyes, 2016](#); [Dehghani et al., 2016](#); [Galarreta and Hestrin, 1998](#); [Graupner and Reyes, 2013](#); [Haider et al., 2006, 2013](#); [Higley and Contreras, 2006](#); [Marino et al., 2005](#); [Okun and Lampl, 2008](#); [Shu et al., 2003](#); [Wehr and Zador, 2003](#); [Wu et al., 2008, 2006](#); [Yizhar et al., 2011](#); [Zhou et al., 2014](#)).

Excitation and inhibition may be balanced in at least two ways. First, inhibitory and excitatory synaptic weights may be co-tuned, so that cells that receive more (or less) excitatory weight receive correspondingly more (or less) inhibitory weight ([Bhatia et al., 2019](#); [Xue et al., 2014](#)). This does not ensure balancing of excitation and inhibition received across varying patterns of activity. Second, given sufficiently strong feedback inhibitory weights, the network dynamics may ensure that the mean inhibition and mean excitation received by neurons remain balanced across patterns of activity, without requiring tuning of synaptic weights. Here, we will focus on this second, dynamic form of balancing.

As we will discuss, theorists have described mechanisms by which inhibition and excitation dynamically remain balanced, keeping neurons in the fluctuation-driven regime, without any need for fine tuning of parameters such as synaptic weights. This dynamical balance can be a “tight balance”, which we define to mean that the excitation and inhibition that cancel are much larger than the residual input that remains after cancellation, or a “loose

balance”, meaning that the canceling inputs are comparable in size to the remaining residual input (terms to describe balanced networks are not yet standardized; see Appendix 1 for comparison of our usage to other nomenclatures). The question of whether the balance is tight or loose has important implications for the behavior of the network. Here we will review these issues and argue that the evidence is most consistent with a loosely balanced regime.

### A Theoretical Problem: How to Achieve Input Mean and Fluctuations That Are Both Comparable in Size to Threshold?

How do cortical neurons stay in the irregularly firing regime? There are two requirements to be in the fluctuation-driven regime, which yields irregular firing: (1) The mean input the neurons receive must be sub- or peri-threshold; (2) Input fluctuations must be sufficiently large to bring neuronal voltages to spiking threshold sufficiently often to create reasonable firing rates. We will measure the voltage effects of a neuron’s inputs in units of the voltage distance from the neuron’s rest to threshold; this distance, typically around 20 mV for a cortical cell (e.g., Constantinople and Bruno, 2013, Fig. 3K), is equal to 1 in these units. Thus, a necessary condition for being in the irregularly firing regime is that the voltage mean driven by the mean input (henceforth abbreviated to “the mean input”) should have order of magnitude 1, which we write as  $\mathcal{O}(1)$ , or smaller. The second requirement above then dictates that the voltage fluctuations driven by the input fluctuations from the mean (henceforth abbreviated to “input fluctuations”) should also be  $\mathcal{O}(1)$ . In particular, this means that the ratio of the mean input to the input fluctuations should be  $\mathcal{O}(1)$ . (Note, we use the  $\mathcal{O}()$  notation simply to indicate order of magnitude, and not in its more technical, asymptotic sense of the scaling with some parameter as that parameter goes to zero or infinity.)

Several authors have considered the requirements for these conditions to be true (Renart et al., 2010; Tsodyks and Sejnowski, 1995; van Vreeswijk and Sompolinsky, 1996, 1998). Following these authors, we assume the network is composed of excitatory ( $E$ ) and inhibitory ( $I$ ) neuron populations, which receive excitatory inputs from an external ( $X$ ) population. The latter could represent any cortical or subcortical neurons outside the local cortical network, for example, the thalamic input to an area of pri-

mary sensory cortex. As a simplified toy model of the assumption that the network is in the asynchronous irregular regime, we assume that the cortical cells fire as Poisson processes without any correlations between them, as do the cells in the external population.

Suppose that a neuron receives  $K_E$  excitatory inputs. Suppose these inputs produce EPSPs that have an exponential time course, with mean amplitude  $J_E$  and time constant  $\tau_E$ , and have mean rate  $r_E$ . Then the mean depolarization produced by these excitatory inputs is  $J_E K_E r_E \tau_E$ . Defining  $n_E = r_E \tau_E$  to be the mean number of spikes of an input in time  $\tau_E$ , we find that the mean excitatory input to the neuron is

$$\mu_E = J_E K_E n_E \quad (1)$$

Let  $\sigma_E$  denote the standard deviation (SD) of fluctuations in this input. Assuming the spike counts of pre-synaptic neurons are uncorrelated, their spike count variances just add. Because they are firing as Poisson processes, the variance in a neuron’s spike count in time  $\tau_E$  is equal to its mean spike count  $n_E$ . Thus, the variance of input from one pre-synaptic neuron is  $J_E^2 n_E$ , and so the variance in the total input is  $K_E J_E^2 n_E$  and

$$\sigma_E = J_E \sqrt{K_E n_E} \quad (2)$$

Therefore the ratio of the mean to the SD of the neuron’s excitatory input is

$$\frac{\mu_E}{\sigma_E} = \sqrt{K_E n_E}, \quad (3)$$

independent of  $J_E$ . Similar reasoning about the neuron’s inhibitory or external input leads to all the same expressions, except with  $E$  subscripts replaced with  $I$  or  $X$  subscripts to represent quantities describing the inhibitory or external input the cell receives.

Again assuming that the different populations are uncorrelated so that their contributed variances add, the *total* or *net* input the neuron receives has mean,  $\mu$ , and standard deviation,  $\sigma$ , given by:

$$\mu = \mu_E + \mu_X - \mu_I \quad (4)$$

$$\sigma = \sqrt{\sigma_E^2 + \sigma_X^2 + \sigma_I^2} \quad (5)$$

We imagine that  $K_E$  and  $K_I$  are the same order of magnitude,  $\mathcal{O}(K)$  for some number  $K$ , and similarly  $n_E$  and  $n_I$  are  $\mathcal{O}(n)$  for some number  $n$ . We also assume  $\mu_X$  and  $\sigma_X$  are the same order of magnitude as  $\mu_E$  or  $\mu_I$  and  $\sigma_E$  or  $\sigma_I$ , respectively, or smaller. Then, if  $\sqrt{Kn}$  is  $\mathcal{O}(1)$ , both  $\mu$  and  $\sigma$  can simultaneously be made  $\mathcal{O}(1)$  with suitable choice of the  $J$ ’s (generically, *i.e.* barring special cases in which

the elements of  $\mu$  precisely cancel). This means that the irregularly-firing regime is self-consistent: having assumed that neurons are in the irregularly-firing regime, we arrive at expressions for the mean and SD of their input that indeed can keep the network in this regime.

If  $K$  is very large, however – large enough that  $\sqrt{Kn} \gg 1$  – then the ratio of the mean to the SD of each type of input, and hence of the net input, is much greater than 1. Van Vreeswijk and Sompolinsky (1996,1998) considered the case of such very large  $K$  and showed how the network could remain in the AI regime. They proposed choosing the  $J$ 's proportional to  $\frac{1}{\sqrt{K}}$ , so that the standard deviations  $\sigma_E$  and  $\sigma_I$  are  $\mathcal{O}(1)$  (Eq. 2), but then by Eq. 3 the means  $\mu_E$  and  $\mu_I$  are large,  $\mathcal{O}(\sqrt{K})$ . Then, for the neurons to be in the asynchronous irregular regime, inhibitory input,  $-\mu_I$ , must cancel or “balance” a sufficient portion of the excitatory input,  $\mu_E + \mu_X$ , so that the mean input,  $\mu$ , is  $\mathcal{O}(1)$ .

If a neuron's mean excitatory and inhibitory inputs very precisely cancel each other, so that the mean net input  $\mu$  is much smaller than either of these factors alone, we say there is *tight balance*. If the net input is more comparable in size to the factors that are cancelling, we call this *loose balance*. The two cases may be distinguished by the size of a dimensionless *balance index*  $\beta$ :

$$\beta = \frac{|\mu|}{\mu_E + \mu_X} \quad (6)$$

Note that, using the above analysis, in the limit of large  $K$  considered by Van Vreeswijk and Sompolinsky (1996,1998),  $\beta \sim \frac{1}{\sqrt{K}}$ . Tight balance means that the balance index is very small,  $\beta \ll 1$ ; in loose balance, the index is not so small, say  $0.1 < \beta < 1$ , very roughly. As we will see, whether the network is in tight or loose balance has important implications for the network's behavior and computational ability. (Note that, in general, the degree of balance can be different in different neurons in the same network. In the above discussion we assumed that different neurons of the same  $E/I$  type are statistically equivalent, e.g., in terms of the number and activity of presynaptic inputs; this is the case in the randomly connected network of Van Vreeswijk and Sompolinsky (1996,1998). In that case  $\beta$  would not vary systematically between neurons of the same type.)

## The Tightly Balanced Solution

Van Vreeswijk and Sompolinsky (1996,1998) showed that, for very large  $K$ , and hence very large  $Kn$ , and all  $J$ 's  $\propto \frac{1}{\sqrt{K}}$ , the network dynamics would produce a tightly balanced solution provided only that some mild (inequality) conditions on the weights are satisfied, without any requirements for fine tuning. This is known more generally as the “balanced network” solution. To understand this solution, we define the mean number of inputs, PSP amplitude, and time constant from population  $B$  ( $B \in \{E, I, X\}$ ) to a neuron in population  $A$  ( $A \in \{E, I\}$ ) to be  $K_{AB}$ ,  $J_{AB}$  and  $\tau_{AB}$  respectively. We define the mean effective weight from population  $B$  to a neuron in population  $A$  as  $W_{AB} = J_{AB}K_{AB}\tau_{AB}$ . Letting  $r_B$  be the average firing rate of population  $B$ , then  $W_{AB}r_B = J_{AB}K_{AB}r_B$ , the mean input from population  $B$  to population  $A$ . We assume that  $W_{AB}r_B = \mathcal{O}(\sqrt{K})$  for all  $A, B$ . The requirements for balance are then that the mean net input to both excitatory and inhibitory cells,  $u_E$  and  $u_I$  respectively, are  $\mathcal{O}(1)$ , where (from Eqs. 1,4),

$$u_E = W_{EE}r_E - W_{EI}r_I + W_{EX}r_X \quad (7)$$

$$u_I = W_{IE}r_E - W_{II}r_I + W_{IX}r_X \quad (8)$$

If we define the external inputs to the network  $l_E = W_{EX}r_X$ ,  $l_I = W_{IX}r_X$ , then these equations can be written as the vector equation

$$\mathbf{u} = \mathbf{W}\mathbf{r} + \mathbf{l} \quad (9)$$

where  $\mathbf{u} \equiv \begin{pmatrix} u_E \\ u_I \end{pmatrix}$ ,  $\mathbf{r} \equiv \begin{pmatrix} r_E \\ r_I \end{pmatrix}$ ,  $\mathbf{l} \equiv \begin{pmatrix} l_E \\ l_I \end{pmatrix}$ , and  $\mathbf{W}$  is the weight matrix  $\mathbf{W} = \begin{pmatrix} W_{EE} & -W_{EI} \\ W_{IE} & -W_{II} \end{pmatrix}$ .

The balanced network solution arises by noting that the left side of Eq. 9 is very small ( $\mathcal{O}(1)$ ) relative to the individual terms on the right ( $\mathcal{O}(\sqrt{K})$ ). So we first find an approximate solution  $\mathbf{r}_0$  to Eq. 9 in which the small left side is replaced by 0 to yield the equation for perfect balance, i.e. all inputs perfectly cancelling:  $\mathbf{W}\mathbf{r}_0 + \mathbf{l} = 0$ , or  $\mathbf{r}_0 = -\mathbf{W}^{-1}\mathbf{l}$ , where  $\mathbf{W}^{-1}$  is the matrix inverse of  $\mathbf{W}$ . Note that  $\mathbf{r}_0$  is  $\mathcal{O}(1)$ , because the elements of  $\mathbf{W}$  and  $\mathbf{l}$  are all the same order of magnitude, so their ratio is generically  $\mathcal{O}(1)$ . We can then write  $\mathbf{r}$  as an expansion in powers of  $\frac{1}{\sqrt{K}}$ ,  $\mathbf{r} = \mathbf{r}_0 + \frac{\mathbf{r}_1}{\sqrt{K}} + \dots$ , where  $\mathbf{r}_0, \mathbf{r}_1, \dots$  are all  $\mathcal{O}(1)$ , to obtain a consistent solution:  $\mathbf{u} = \frac{\mathbf{W}\mathbf{r}_1}{\sqrt{K}} + \dots$  where the first term on the right is  $\mathcal{O}(1)$ , as desired, and the remaining terms are very small ( $\mathcal{O}\left(\frac{1}{\sqrt{K}}\right)$  or smaller).

The authors showed that, with some mild general conditions on the weights  $\mathbf{W}$  and inputs  $\mathbf{I}$ , this tightly balanced solution would be the unique stable solution of the network dynamics. That is, for a given fixed input  $\mathbf{I}$ , the network's excitatory/inhibitory dynamics will lead it to flow to this balanced solution for the mean rates:  $\mathbf{r} = -\mathbf{W}^{-1}\mathbf{I} + \mathcal{O}\left(\frac{1}{\sqrt{K}}\right)$ .

We immediately see two points about the tightly balanced ( $\sqrt{Kn} \gg 1$ ) solution:

1. *Mean population responses are linear in the inputs.*  $-\mathbf{W}^{-1}\mathbf{I}$  is a linear function of the input  $\mathbf{I}$ . Tight balance implies that nonlinear corrections to  $\mathbf{r} \approx \mathbf{r}_0 = -\mathbf{W}^{-1}\mathbf{I}$  are very small relative to this linear term, except for very small external inputs, so mean response  $\mathbf{r}$  is for practical purposes a linear function of the input.
2. *External input must be large relative to the net input and to the distance from rest to threshold.* The external input  $\mathbf{I}$  must have the same order of magnitude as the recurrent input  $\mathbf{W}\mathbf{r}_0$ , so that balance can occur,  $\mathbf{W}\mathbf{r}_0 = -\mathbf{I}$ , with rates that are  $\mathcal{O}(1)$ . If  $\mathbf{I}$  were smaller, the firing rates  $\mathbf{r} \approx \mathbf{r}_0$  would correspondingly be unrealistically small.

In the above treatment we focused on population-averaged responses,  $r_E$  and  $r_I$ . We emphasize that the balancing only applies to the mean input across neurons of each type, and leaves unaffected input components with mean of zero across a given type; while the mean input is very large in the tightly balanced regime, zero-mean input components can be  $\mathcal{O}(1)$  and yet elicit  $\mathcal{O}(1)$  responses in individual neurons (see e.g. (Hansel and van Vreeswijk, 2012; Pehlevan and Sompolinsky, 2014; Sadeh and Rotter, 2015)). Furthermore, even in the tightly balanced regime, individual neurons can exhibit nonlinearities in their responses, but these are washed out at the level of population-averaged responses. We also note that synaptic nonlinearities, e.g. synaptic depression, which were neglected here, can allow a tightly balanced state with nonlinear population-averaged responses (Mongillo et al., 2012).

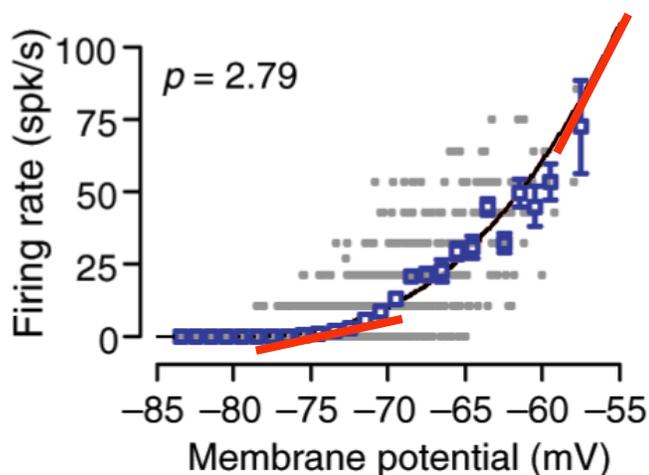
## A Loosely Balanced Regime

As we have seen, if  $Kn$  is  $\mathcal{O}(1)$ , the mean and the fluctuations of the input that neurons receive can both be

$\mathcal{O}(1)$  without requiring any balancing. Given that there is both excitatory and inhibitory input, there will always be some input cancellation or "balancing" – some portion of the input excitation will be cancelled by input inhibition, leaving some smaller net input. When  $Kn$  is  $\mathcal{O}(1)$ , all of these quantities – the excitatory input, the inhibitory input, and the net input after cancellation – will generically be  $\mathcal{O}(1)$ , and thus balancing is "loose" – the factors that cancel and the net input after cancellation are of comparable size, and the balance index  $\beta$  is not small.

However, the fact that there is some inhibition that cancels some excitation does not by itself imply interesting consequences for network behavior. We will use the term "loosely balanced solution" to refer more specifically to a solution having two features: (1) the dynamics yields a systematic cancellation of excitation by inhibition like that in the tightly balanced solution. In particular, in the loosely balanced networks on which we will focus, a signature of this systematic cancellation is that the net input a neuron receives grows sublinearly as a function of its external input (we will make this more precise below); (2) this cancellation is "loose", as just described. As we will discuss, such a loosely balanced solution produces various specific nonlinear network behaviors that are observed in cortex.

Ahmadian et al. (2013) showed that such a loosely balanced solution would naturally arise from E/I dynamics for recurrent weights and external inputs that are not large, provided that the neuronal input/output function, determining firing rate vs. input level, is supralinear (having ever-increasing slope) over the neuron's dynamic range. They modeled this supralinear input/output function as a power law with power greater than 1 (Fig. 1). Such a power-law input-output function is theoretically expected for a spiking neuron when firing is induced by input fluctuations rather than the input mean (Hansel and van Vreeswijk, 2002; Miller and Troyer, 2002), and is observed in intracellular recordings over the full dynamic range of neurons in primary visual cortex (V1) (Priebe and Ferster, 2008). Of course, a neuron's input/output function must ultimately saturate but, at least in V1, the neurons do not reach the saturating portion of their input/output function under normal operation. For the loosely balanced solution to arise, some general conditions on the weight matrix, similar to those for the tightly balanced network solution but less restrictive, must also be satisfied.



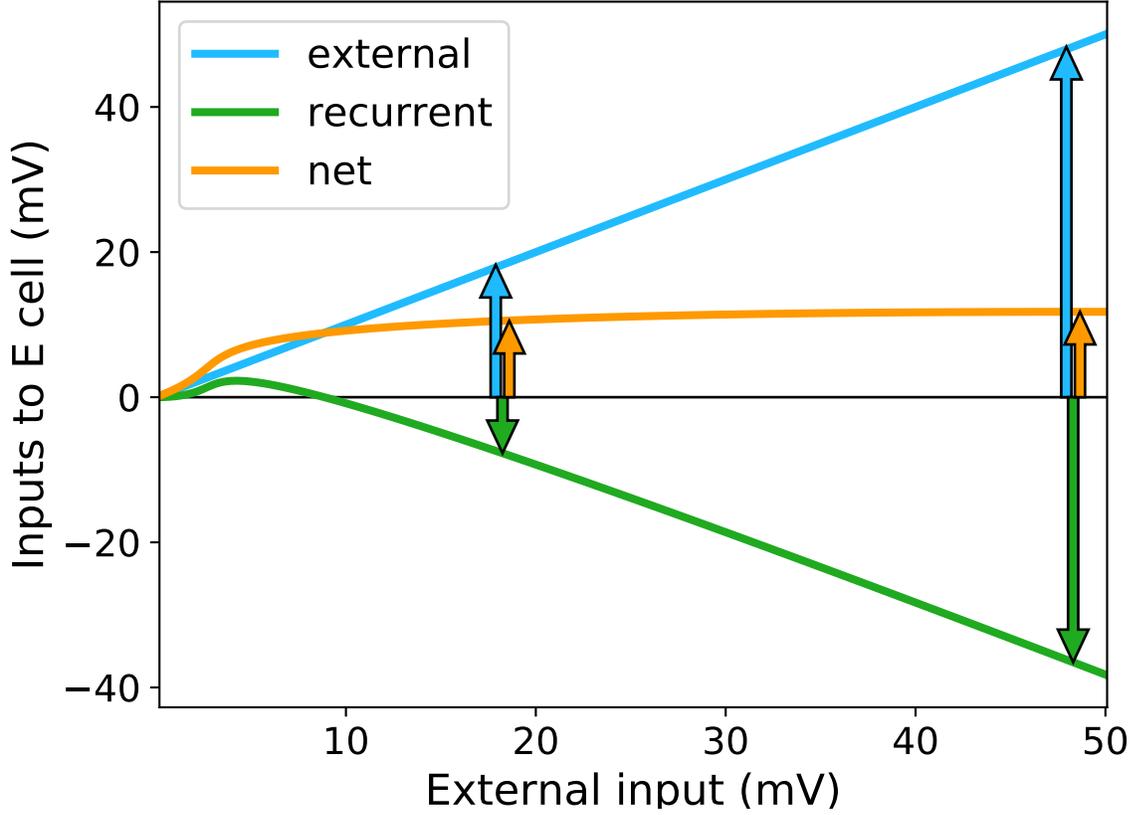
**Figure 1. The supralinear (power-law) neuronal transfer function.** The transfer function of neurons in cat V1 is non-saturating in the natural dynamic range of their inputs and outputs, and is well fit by a supralinear rectified power-law with exponents empirically found to be in the range 2-5. Such a curve exhibits increasing input-output gain (i.e. slope, indicated by red lines) with growing inputs, or equivalently with increasing output firing rates. Gray points indicate a studied neuron’s average membrane potential and firing rate in 30ms bins; blue points are averages for different voltage bins; and black line is fit of power law,  $r = [V - \theta]_+^p$ , where  $r$  is firing rate,  $V$  is voltage,  $[x]_+ = x, x > 0, =0$  otherwise;  $\theta$  is a fitted threshold; and  $p$ , the fitted exponent, here is 2.79. Figure modified from (Priebe et al., 2004).

In the presence of a supralinear input/output function, the loosely balanced solution arises as follows. Whereas previously we considered the effects of increasing  $K$  when recurrent and external inputs were all  $\mathcal{O}(\sqrt{K})$ , now we consider the more biological case of increasing external input (i.e., stimulus) strength while recurrent weights are at some fixed level. The supralinear input/output function means that a neuron’s gain – its change in output for a given change in input – is continuously increasing with its activation level. This in turn means that effective synaptic strengths are increasing with increasing network activation. The effective synaptic strength measures the change in the postsynaptic cell’s firing rate per change in presynaptic firing rate. This is the product of the actual synaptic strength – the change in postsynaptic input induced by a change in presynaptic firing – and the

postsynaptic neuron’s gain, hence the effective synaptic strengths increase with increasing gains.

The increasing effective synaptic strengths lead to two regimes of network operation. For very weak external drive and thus weak network activation, all effective synaptic strengths are very weak, for both externally-driven and network-driven synapses. External drive to a neuron is delivered monosynaptically, via the weak externally-driven synapses. In contrast, assuming that the network is inactive in the absence of external input, network drive involves a chain of two or more weak synapses: the weak externally driven synapses activate cortical cells, which then drive the weak network-driven synapses. From the same principle that  $x^2 \ll x$  when  $x \ll 1$ , the network drive is therefore much weaker than the external drive. Thus, the input to neurons is dominated by the external input, with only relatively small contributions from recurrent network input. In sum, in this weakly-activated regime, the neurons are *weakly coupled*, largely responding directly to their external input with little modification by the local network.

With increasing external (stimulus) drive and thus increasing network activation, the gains and thus the effective synaptic strengths grow. This causes the relative contribution of network drive to grow until the network drive is the dominant input. At some point, the effective  $E \rightarrow E$  connections become strong enough that the network would be prone to instability – a small upward fluctuation of excitatory activities would recruit sufficient recurrent excitation to drive excitatory rates still higher, which if unchecked would lead to runaway, epileptic activity (and to ever-growing effective synaptic strengths and thus ever-more-powerful instability). However, if feedback inhibition is strong and fast enough, the inhibition will stabilize the network, that is, it becomes an ISN. This stabilization is achieved by a loose balancing of excitation and inhibition, as we will explain in more detail below. Thus, in this more strongly-activated regime, the neurons are *strongly coupled* and are *loosely balanced*. Note that the input driving spontaneous activity may already be strong enough to obscure the weakly coupled regime, as suggested by the finding that V1 under spontaneous activity is already an ISN (Sanzeni et al., 2019). As in the tightly balanced network, the network’s excitatory/inhibitory dynamics lead it to flow to this loosely balanced solution, without any need for fine tuning of parameters. Because this mechanism involves stabilization, by inhibition, of the instability induced by the supralinear



**Figure 2. Loose vs. tight balance.** We illustrate the external (blue), recurrent or within-network (green) and net (orange, equal to external plus recurrent) inputs to a typical excitatory cell in a Stabilized Supralinear Network. At all biological ranges of external input (stimulus) strength, the balance is loose, as exhibited by the left set of three arrows (representing the external, recurrent and net inputs): the net input is comparable in size to the other two. Nevertheless the balance systematically tightens with increasing external input (right set of arrows), as the net input grows only sublinearly with growing external input strength. At very high (possibly non-biological) levels of external input, the balance can become very tight, with the net input much smaller in magnitude than the external and recurrent inputs.

input/output function of individual neurons along with  $E \rightarrow E$  connections, it has been termed the *Stabilized Supralinear Network* (SSN) (Ahmadian et al., 2013; Rubin et al., 2015).

To describe the mathematics of this mechanism, we again consider an excitatory and an inhibitory population along with external input. We define the vectors  $\mathbf{r}$ ,  $\mathbf{u}$  and  $\mathbf{I}$  and the matrix  $\mathbf{W}$  as before. Then the power-law input/output function means that the network's steady state firing rate  $r_{SS}$  for a steady input  $\mathbf{I}$  satisfies

$$r_{SS} = k(\mathbf{u})_+^p = k(\mathbf{W}\mathbf{r}_{SS} + \mathbf{I})_+^p \quad (10)$$

where  $(\mathbf{v})_+$  is the vector  $\mathbf{v}$  with negative elements set to zero,  $(\mathbf{v})_+^p$  means raising each element of  $(\mathbf{v})_+$  to the power  $p$ ,  $p$  is a number greater than 1 (typically, 2 to 5, Priebe and Ferster, 2008), and  $k$  is a constant with units  $\frac{\text{Hz}}{(\text{mV})^p}$  (and the units of  $\mathbf{W}$ ,  $\mathbf{r}$ , and  $\mathbf{I}$  are  $\frac{\text{mV}}{\text{Hz}}$ ,  $\text{Hz}$ , and  $\text{mV}$  respectively). It is convenient to absorb  $k$  into effective weights and inputs by writing  $\tilde{\mathbf{W}} = k^{1/p}\mathbf{W}$ ,  $\tilde{\mathbf{I}} = k^{1/p}\mathbf{I}$ , so the equation becomes

$$r_{SS} = (\tilde{\mathbf{W}}\mathbf{r}_{SS} + \tilde{\mathbf{I}})_+^p \quad (11)$$

If we let  $\psi = \|\tilde{\mathbf{W}}\|$  represent a norm of  $\tilde{\mathbf{W}}$  (think of it as a typical total E or I recurrent weight received by a neuron), and similarly let  $c = \|\tilde{\mathbf{I}}\|$  represent a typical in-

put strength, then it turns out the network regime is controlled by the dimensionless parameter

$$\alpha = \psi c^{p-1}. \quad (12)$$

As we increase the strength of external drive and thus of network activation,  $c$  and thus  $\alpha$  is increasing. When  $\alpha \ll 1$ , the network is in the weakly coupled regime; for  $\alpha \gg 1$ , the network is in the strongly coupled regime; and the transition between regimes generically occurs when  $\alpha$  is  $\mathcal{O}(1)$  (Ahmadian et al., 2013).

The loosely-balanced solution then turns out to be of the form

$$\mathbf{r} = -\mathbf{W}^{-1}\mathbf{I} + \frac{c}{\psi} \left( \frac{\mathbf{r}_1}{\alpha^{1/p}} + \dots \right) \quad (13)$$

where  $\mathbf{r}_1$  is dimensionless and  $\mathcal{O}(1)$ , and the higher-order terms (indicated by ...) involve higher powers of  $\frac{1}{\alpha^{1/p}}$  (see Appendix 2). Equation 13 is precisely the same equation as for the tightly balanced solution, in the case that the input/output function is a power law. In the tightly balanced network,  $\psi$  and  $c$  are both  $\mathcal{O}(\sqrt{K})$ , so  $\alpha$  is  $\mathcal{O}((K)^{p/2})$ , i.e. very large, and the  $\frac{1}{\alpha^{1/p}}$  in the second term becomes  $\mathcal{O}\left(\frac{1}{\sqrt{K}}\right)$ , as expected. The loosely-balanced solution arises, however, when  $\alpha$  is  $\mathcal{O}(1)$ . In particular, in the biological case of fixed weights but increasing stimulus drive, and given the supralinear neuronal input/output functions, the same E/I dynamics that lead to the tightly balanced solution when inputs are very large will already yield a loosely balanced solution when inputs are  $\mathcal{O}(1)$ . The conditions for this loosely balanced solution to arise are further discussed in Appendix 2.

The fact that the solution is loosely balanced can be seen by computing the balance index,  $\beta$  (Eq. 6). The network excitatory drive is  $\mathcal{O}(\psi\mathbf{r})$ , the external drive is  $\mathcal{O}(c)$ , and because the first term on the right side of Eq. 13 cancels the external input, the net input after balancing is  $\mathbf{W}$  (which is  $\mathcal{O}(\psi)$ ) times the 2nd term on the right side of Eq. 13, or  $\mathcal{O}\left(\frac{c}{\alpha^{1/p}}\right) = \mathcal{O}\left(\left(\frac{c}{\psi}\right)^{1/p}\right)$ . Since  $1/p < 1$ , the net input thus grows sublinearly with growing external input strength,  $c$ , as illustrated in Fig. 2. Moreover, it follows that the balance index (Eq. 6) is  $\mathcal{O}\left(\frac{c/\alpha^{1/p}}{c+\psi\mathbf{r}}\right)$  which is  $\mathcal{O}\left(\frac{1}{\alpha^{1/p}}\right)$  (assuming the order of magnitude of the recurrent input,  $\psi\mathbf{r}$ , is the same as or smaller than that of the external input strength,  $c$ ). Again, for the tightly balanced solution this is very small,  $\mathcal{O}\left(\frac{1}{\sqrt{K}}\right)$ , but the loosely balanced solution arises when this is  $\mathcal{O}(1)$ .

In more complex models (involving many neurons with structured connectivity and stimulus selectivity), loosely-balanced solutions still arise when  $\alpha$  is  $\mathcal{O}(1)$ . That is, even in such cases the full nonlinear steady state equations, Eq. (10), can yield biologically plausible solutions, and when that happens the net inputs to activated neurons grow sublinearly with growing external input strength, and balance indices are  $\mathcal{O}(1)$ . The case of structured networks with stimulus selectivity is further discussed in Appendix 2.

We can now see that the loosely balanced regime differs from the tightly balanced in the two points summarized previously:

1. *In the loosely balanced regime, mean population responses are nonlinear in the inputs.* This is because, when balance is loose, the second term in Eq. 13, which is not linear in the input, cannot be neglected relative to the first, linear term. In particular, the nonlinear population behaviors observed in the loosely balanced regime with a supralinear input/output function closely match the specific nonlinear behaviors observed in sensory cortex (Rubin et al., 2015), as we will discuss below.
2. *In the loosely balanced regime, external input can be comparable to the net input and to the distance from rest to threshold.*

## What Regime Do Experimental Measurements Suggest?

As we have seen above, the same model can give a loosely balanced solution (Eq. 13) when  $\alpha$  is  $\mathcal{O}(1)$  (e.g., when  $c$  and  $\psi$  are both  $\mathcal{O}(1)$ ), but give a tightly balanced solution when  $\alpha$  is large (e.g., when  $c$  and  $\psi$  are both  $\mathcal{O}(\sqrt{K})$ ). Which of these regimes is supported by experimental measurements?

## Measurements of Biological Parameters

**How large is  $\sqrt{K\eta}$ ?** We saw in Eq. 3 that the ratio of the mean to the standard deviation,  $\mu_Y/\sigma_Y$ , of the input of type  $Y$  ( $Y \in \{E, I, X\}$ ) received by a neuron is equal to  $\sqrt{K_Y n_Y}$ , where  $K_Y$  is the number of inputs of type  $Y$  a given neuron receives and  $n_Y$  is the average number of spikes one of these inputs fires in a PSP decay time  $\tau_Y$

	$K_E = 200$	$K_E = 1000$	$K_E = 5000$
$r_E = 0.1$ Hz	0.6	1.4	3.2
$r_E = 1$ Hz	2.0	4.5	10.0
$r_E = 10$ Hz	6.3	14.1	31.6

**Table 1.** Values of  $\sqrt{K_E n_E}$  for varying  $K_E$  and  $r_E$ , for  $\tau = 20$ ms.

( $n_Y = r_Y \tau_Y$ , where  $r_Y$  is the average firing rate of one of these inputs). Here we estimate  $\sqrt{K_E n_E}$ .

Note that  $\sqrt{K_Y n_Y}$  is actually an upper bound for the ratio  $\mu_Y/\sigma_Y$  for a given input type, because we have neglected a number of factors that would increase fluctuations for a given mean. These include (i) correlations among inputs which, even if weak, can significantly boost the input SD,  $\sigma_Y$ , without altering input mean; (ii) variance in the weights,  $J_Y$ , which would increase the estimate of  $\sigma_Y$  by a factor  $\sqrt{\langle J_Y^2 \rangle / \langle J_Y \rangle^2}$ ; and (iii) network effects that can amplify input variances by creating firing rate fluctuations, although this amplification may be small for strong stimuli (Hennequin et al., 2018). Furthermore, the ratio  $\mu/\sigma$  of total input is expected to be smaller than the ratio  $\mu_Y/\sigma_Y$  for any single type. This is because  $\sigma^2$  involves the sum of three variances (Eq. 5), while  $\mu$  involves a difference of one mean from the sum of two others (Eq. 4), representing the effect of loose balancing.

Given these considerations, we are primarily concerned with estimating the overall magnitude of  $\sqrt{K_E n_E}$  rather than detailed values. If this magnitude is very much larger than the observed  $\mu/\sigma$  ratio *in vivo*, then tight balancing may be needed to explain the *in vivo* ratio. To estimate the *in vivo*  $\mu/\sigma$  ratio, we note that, in anesthetized cat V1,  $\sigma$  varies from 1 to 7 mV and  $\mu$  ranges from 0 to 15 mV (20 mV in one case) for a strong stimulus (Finn et al., 2007; Sadagopan and Ferster, 2012). While these authors did not give the paired  $\mu$  and  $\sigma$  values for individual cells, it seems reasonable to guess from these values that the value of  $\mu/\sigma$  for the total input to these cells is generally in the range 0–15. Finn et al. (2007) also reported that, at the peak of a simple cell’s voltage modulation to a high-contrast drifting grating, the ratio  $\sigma/\mu$  had an average value of about 0.15 (here, we are taking  $\mu$  to be the mean voltage at the peak). This suggests that the average value of  $\mu/\sigma$  at peak activation is around  $1/0.15 = 6.7$ .

How large is  $\sqrt{K_E n_E}$ ? In a study of input to excitatory cells in layer 4 of rat S1 (Schoonover et al., 2014), the EPSP decay time  $\tau_E$  was around 20 ms. From 1800 to 4000 non-thalamic-recipient spines were found on stud-

ied cells which, with an estimated average of 3.4 synapses per connection between layer 4 cells (Feldmeyer et al., 1999), corresponds to a  $K_E$  – the number of other cortical cells providing input to one cell – of 530 to 1200. If  $r_E$  is expressed in Hz, then  $\sqrt{K_E n_E}$  ranges from  $3.3\sqrt{r_E}$  to  $4.9\sqrt{r_E}$ . Thus, even if average input firing rates were 10 Hz, which would be very high for rodent S1 (Barth and Poulet, 2012) (note that the average is over all inputs, not just those that are well driven in a given situation, and so is likely far below the rate of a well-driven neuron), these ratios would be 10.4–15.5. For more realistic rates of 0.1–1 Hz (Barth and Poulet, 2012), these ratios would be 1.0–4.9. All of these are comparable in magnitude to observed *in vivo* levels of  $\mu/\sigma$ .

More generally, estimates across species and cortical areas of the number of spines on excitatory cells, and thus of the number of excitatory synapses they receive, range from 700 to 16,000, with numbers increasing from primary sensory to higher sensory to frontal cortices (Amatrudo et al., 2012; Elston, 2003; Elston and Fujita, 2014; Elston and Manger, 2014). Estimates of the mean number of synapses per connection between excitatory cells range from 3.4 to 5.5 across different areas and layers studied (Fares and Stepanyants, 2009; Feldmeyer et al., 1999, 2002; Markram et al., 1997). These numbers yield a  $K_E$  of 130 to 4700. In Table 1, we show the value of  $\sqrt{K_E n_E}$  for  $K_E$  ranging from 200 to 5000 (rounded upward to bias results most in favor of a need for tight balancing) and for rates  $r_E$  of 0.1 to 10 Hz. The results are all comparable to the  $\mu/\sigma$ ’s observed *in vivo*, except for the most extreme case considered ( $K_E = 5000$ ,  $r_E = 10$ Hz), and even that case is only off by a factor of 2. Thus, the numbers strongly argue that tight balancing is not needed for the ratio of voltage mean to variance to have values as observed *in vivo*.

**External input is comparable in strength to net input.** Several studies have silenced cortical firing while recording intracellularly to determine the strength of external input, with cortex silenced, relative to the net input with cortex intact. These find the external input to

be comparable to the net input, consistent with the loose balance scenario, rather than much larger as the tight balance scenario requires.

Ferster et al. (1996) cooled V1 and surrounding V2 in anesthetized cats to block spiking of almost all cortical cells, both excitatory and inhibitory, leaving axonal transmission (e.g. of thalamic inputs) intact, though with weakened release. By measuring the size of EPSPs evoked by electrical stimulation of the thalamic lateral geniculate nucleus (LGN) in thalamic-recipient cells in layer 4 of V1, they could estimate the degree of voltage attenuation of EPSPs induced by cooling. Correcting for this attenuation, they estimated that the first harmonic voltage response to an optimal drifting luminance grating stimulus of a layer 4 V1 cell was, on average, about 1/3 as great with cortex cooled as with cortex intact, suggesting that the external input to cortex is *smaller* than the net input with cortex intact. Chung and Ferster (1998) and Finn et al. (2007) assayed the same question by using cortical shock to silence the local cortex for about 150 ms, during which time the voltage response to an optimal flashed grating was measured. They found that on average the transient voltage response in layer 4 cells with cortex silenced was about 1/2 the size of that with cortex intact (Chung and Ferster, 1998), and more generally ranged from 0% to 100% of the intact cortical response (Finn et al., 2007). This again suggests that the external input to cortex is *smaller* than the net input, *i.e.* the external input is  $\mathcal{O}(1)$ , consistent with loose but not tight balance.

**Total excitatory or inhibitory conductance is comparable to threshold.** The above results suggest that depolarization due to thalamus alone is less than that induced by the combination of thalamic and cortical excitation plus cortical inhibition, *i.e.* after cortical "balancing" has occurred. One can also ask what proportion of the total excitation is provided by thalamus. This has been addressed in voltage-clamp recordings in anesthetized mice by silencing cortex through light activation of parvalbumin-expressing inhibitory cells expressing channelrhodopsin. In layer 4 cells of V1 (Li et al., 2013b; Lien and Scanziani, 2013) and primary auditory cortex (A1) (Li et al., 2013a), mean stimulus-evoked excitatory conductance with cortex silenced was estimated to be 30-40% of that with cortex intact.

This tells us that the external and cortical contributions to excitation are comparable. How large are they compared to the excitation needed to depolarize the cell from

rest to threshold, which is typically a distance of about 20 mV (Constantinople and Bruno, 2013)? With cortical spiking intact, these authors (Li et al., 2013a; Lien and Scanziani, 2013) found mean stimulus-evoked peak excitatory currents ranging from 60 to 150pA for various stimuli. Even assuming a membrane resistance of 200 M $\Omega$ , which seems on the high end for *in vivo* recordings (Li et al. (2013a) reported input resistances of 150 – 200 M $\Omega$ ), these would induce depolarizations of 12 to 30 mV; that is, the total excitatory current is comparable to threshold, *i.e.* it is  $\mathcal{O}(1)$ .

A similar result can be found from decomposition of excitatory and inhibitory conductances from current-clamp recordings at varying hyperpolarizing current levels. In cat V1 cells for an optimal visual stimulus, one finds that peak excitatory and inhibitory stimulus-induced conductances,  $g_E$  and  $g_I$  respectively, are typically  $< 10$ nS and almost always  $< 20$ nS, on top of stimulus-independent conductances ( $g_L$ , for leak conductance) around 10nS (Anderson et al., 2000; Ozeki et al., 2009). A study of response to whisker stimulation in rat barrel cortex found excitatory and inhibitory conductances of  $\leq 5$ nS (Lankarany et al., 2016). The depolarization that the stimulus-induced excitatory conductance would induce by itself is  $\frac{g_E}{g_E+g_L} V_E$ , where  $V_E$  is the driving potential of excitatory conductance, about 50 mV at spike threshold of around  $-50$  mV (e.g. Wilent and Contreras, 2005). Using the cat V1 numbers, this means that the depolarization driven by excitatory conductance is typically  $< 25$  mV and almost always  $< 33$  mV. Hyperpolarization driven by the inhibitory conductance alone would be 0.4 to 0.6 times these values, given inhibitory driving force of  $-20$  to  $-30$  mV at spike threshold. These values are all quite comparable to the distance from rest to threshold,  $\sim 20$  mV, that is, they are  $\mathcal{O}(1)$ .

### How large is the expected mean excitatory input?

We have seen that the expected mean depolarization induced by recurrent excitation is  $J_E K_E n_E$ , where  $J_E$  is the mean EPSP amplitude. Based on the measurements of Lien and Scanziani (2013) and Li et al. (2013b), discussed above, total excitation may be about 1.5 times greater than recurrent excitation.  $J_E$  can be difficult to estimate, because some of the  $K_E$  anatomical synapses may be very weak and not sampled in physiology, and because synaptic failures, depression, or facilitation can alter average EPSP size relative to measured EPSP sizes. Furthermore, measurements are variable, for example  $J_E$  for layer 4 to layer 4 connections in rodent barrel cortex has been es-

estimated to be 1.6 mV *in vitro* (Feldmeyer et al., 1999) vs. 0.66 mV *in vivo* (Schoonover et al., 2014). If we assume typical values for  $J_E$  of 0.5 – 1 mV, then  $1.5J_E K_E n_E$  would exceed 75 mV for  $\sqrt{Kn} > 7 - 10$  and exceed 150 mV for  $\sqrt{Kn} > 10 - 14$  (compare values of  $\sqrt{Kn}$  in Table 1). We can very roughly guess that neural responses may become better described by tight rather than loose balance somewhere in this range of mean excitatory input (and corresponding  $\sqrt{Kn}$ ). While the measurements of excitatory currents and conductances described above argue that such a range is not reached in primary sensory cortex, it could conceivably be reached (Table 1) in areas with higher  $K_E$ , e.g. frontal cortex.

## Nonlinear Behaviors

Sensory cortical neuronal responses display a variety of nonlinear behaviors that, as we'll describe, are expected from the SSN loosely balanced regime but not from the tightly balanced regime. Many of these nonlinearities are often summarized as "normalization" (Carandini and Heeger, 2012), meaning that responses can be fit by a phenomenological model of an unnormalized response that is divided by (normalized by) some function of all of the unnormalized responses of all the neurons within some region. To describe these nonlinear behaviors, we must first define the classical receptive field (CRF): the localized region of sensory space in which appropriate stimuli can drive a neuron's response.

One nonlinear property is sublinear response summation: across many cortical areas, the response to two stimuli simultaneously presented in the CRF is less than the sum of the responses to the individual stimuli, and is often closer to the average than the sum of the individual responses (reviewed in Carandini and Heeger, 2012; Reynolds and Chelazzi, 2004). An additional nonlinearity is that the form of the summation changes with the strength of the stimulus: summation becomes linear for weaker stimuli (Heuer and Britten, 2002; Ohshiro et al., 2013). It is often difficult to determine if such nonlinear behaviors are computed in the recorded area or involve changes in the inputs to that area. However, some recent experiments studied summation of response to an optogenetic and a visual stimulus, a case in which the inputs driven by each stimulus should not alter those driven by the other. Sublinear summation of responses to these stimuli was found (Nassi et al., 2015; Wang et al., 2019, but see Histed, 2018), which became linear for weak stimuli (Wang et al., 2019).

Another set of nonlinearities involve interaction of a CRF stimulus and a "surround" stimulus, which is located outside the CRF. Across many cortical areas, surround stimuli can suppress response to a CRF stimulus ("surround suppression"; reviewed in Angelucci et al., 2017; Rubin et al., 2015), but this effect varies with stimulus strength. When the center stimulus is weak, a surround stimulus can facilitate rather than suppress response (Ichida et al., 2007; Polat et al., 1998; Sato et al., 2014; Schwabe et al., 2010; Sengpiel et al., 1997). Similarly, the *summation field size* – the size of a stimulus that elicits strongest response, before further expansion yields surround suppression – is largest for weak stimuli and shrinks with increasing stimulus strength (Anderson et al., 2001; Cavanaugh et al., 2002; Nienborg et al., 2013; Sceniak et al., 1999; Shushruth et al., 2009; Song and Li, 2008; Tsui and Pack, 2011). The summation field size in feature space – the optimal range of simultaneously presented motion directions in monkey area MT – similarly shrinks with increasing stimulus strength (Liu et al., 2018).

Additional nonlinearities include a decrease, with increasing stimulus strength, in the ratio of excitation to inhibition received by neurons (Adesnik, 2017) and in the wavelength of a characteristic spatial oscillation of activity (Rubin et al., 2015).

All of these nonlinear cortical response properties, and more, follow naturally (Ahmadian et al., 2013; Rubin et al., 2015) from the two regimes of the loosely balanced scenario with a supralinear input/output function, along with simple assumptions on connectivity (e.g. that connections decrease in strength and/or probability with spatial distance, e.g. Markov et al., 2011, or with difference in preferred features, e.g. Cossell et al., 2015; Ko et al., 2011). In contrast, as described previously, the tightly balanced scenario causes population-averaged responses to be linear responses to the input (individual neurons, but not the population average, may have nonlinear behaviors), and thus appears inconsistent with these nonlinear cortical behaviors, which in most cases are consistent enough across neurons that they should characterize the mean population response. Synaptic nonlinearities can give nonlinear population-averaged behavior in the tightly balanced regime (Mongillo et al., 2012), but it has not been claimed or demonstrated that this could produce specific nonlinearities like those seen in cortical responses.

## Correlations and Variability

Across many cortical systems, the correlated component of neuronal variability is decreased when a stimulus is given, with variability decrease seen both in neurons that respond to the stimulus and those that don't respond (Churchland et al., 2010). This is also naturally explained by the loosely balanced SSN network (Hennequin et al., 2018). In the strongly coupled regime of the loosely balanced SSN network, increasing stimulus strength increases the strength with which correlated patterns of activity inhibit themselves, thus damping their responses to input fluctuations. The tightly balanced state represents the end state of this process – a fully asynchronous regime in which correlations are completely suppressed (Renart et al., 2010; van Vreeswijk and Sompolinsky, 1998; with dense connectivity, the mean correlation is proportional to  $1/K$ , and the standard deviation of the distribution of correlations is proportional to  $1/\sqrt{K}$  Renart et al. (2010); recall that  $K$  is meant to be a very large number to achieve the tightly balanced state). Thus, the tightly balanced state appears incompatible with the observed decrease in correlated variability induced by a stimulus, because the state has essentially no correlated variability. However, it should be noted that variants of the tightly balanced network involving structured connectivity can yield finite correlated variability among preferentially connected neurons while maintaining tight balance, although average correlation over all neuron pairs can still go to zero with increasing  $\sqrt{K}$  (Litwin-Kumar et al., 2012; Rosenbaum et al., 2017).

## Discussion

We have seen that many independent lines of evidence are all consistent with cortex being in a loosely balanced regime, and are inconsistent with tight balance. We define balance to mean that the dynamics yields a systematic cancellation of excitation by inhibition; a signature of this for the loosely balanced scenario that we consider is that the net input a neuron receives after cancellation grows sublinearly as a function of its external input. Loose balance means that the net input after cancellation is comparable in size to the factors that cancel, whereas tight balance means that the net input is very small relative to the cancelling factors. In both cases, the net input after cancellation is comparable in size to the distance from rest to threshold so that neuronal firing can

be in the fluctuation-driven regime that produces irregular firing like that observed in cortex.

One line of evidence for loose balance involves a variety of measurements on the numbers and/or strengths of the inputs cells receive, including spine counts, strengths of external and total input, and strengths of excitatory and of inhibitory input. These measurements show that the expected ratio of mean to standard deviation of the network input before any tight balancing is already consistent with the ratios observed for a cell's net input as judged by its voltage response. That is, tight cancellation is not needed to achieve the ratios observed. These measurements further show that external input and network input are comparable in size to the net input remaining after cancellation, and that they and the total excitatory and total inhibitory input are all comparable to the distance from rest to threshold, consistent with loose but not tight balancing. Other lines of evidence include a variety of nonlinear population response properties of sensory cortical neurons, as well as the presence of correlated variability in neural responses and its decrease upon presentation of a stimulus, all of which emerge naturally from loose balance with a supralinear input/output function, but appear largely incompatible with tight balance.

It should be emphasized that the number of excitatory synapses received by an excitatory cell,  $K_E$ , increases from primary sensory to higher sensory to frontal cortex (e.g. Elston, 2003). Higher numbers are expected to push in the direction of tighter balance. The expected ratio of input mean to standard deviation and the expected size of the mean input both can become high enough to potentially yield tight balance for the highest  $K_E$ 's, particularly if higher average firing rates  $r_E$  are assumed. Our other arguments depend largely, but not entirely, on measurements from sensory cortex. The measurements of net input and external input are all from primary sensory cortex. The studied nonlinear response properties are primarily from both lower and higher visual cortices (reviewed in Rubin et al., 2015). Suppression of correlated variability by a stimulus, however, has been observed in frontal and parietal as well as sensory cortex (Churchland et al., 2010). In sum, while the evidence strongly favors loose balance in sensory cortex, the evidence as to the regime of frontal cortex is weaker.

The seminal discovery of the tightly balanced network (van Vreeswijk and Sompolinsky, 1996, 1998) solved a key problem in theoretical neuroscience: how can neurons re-

main in the fluctuation-driven regime, so that they have irregular firing with reasonable firing rates, without requiring fine tuning of parameters? The answer was that when external and network inputs were very large, the network’s dynamics could robustly tightly balance the excitation and inhibition that neurons receive, leaving a net input after cancellation that is negligibly small relative to the factors that cancel. This allows both the mean and standard deviation of the net input to be comparable to the distance from rest to threshold despite the very large size assumed for the factors that cancel, yielding the fluctuation-driven regime. This achievement along with the model’s mathematical tractability have made it a popular model for the theoretical study of neural circuits. However, for all of the reasons stated above, this tightly balanced regime does not seem to match observations of at least sensory cortical anatomy and physiology.

The loosely balanced solution shows that, when neuronal input/output functions are supralinear, the same dynamical balancing can arise from network dynamics without fine tuning, but in a regime in which external and network inputs are not large. Instead, the balancing arises when these inputs, and the net input remaining after cancellation of excitatory and inhibitory input, are all comparable in size to one another and to the distance from rest to threshold. Furthermore, for weak inputs this same scenario produces a weakly-coupled, feedforward-driven regime which explains the observation that summation changes from sublinear, or suppressive, for stronger stimuli to linear, or facilitative, for weak inputs.

The tightly balanced network demonstrated that a network could self-consistently generate its own variability. As we showed in the section “How large is  $\sqrt{Kn}$ ?”, the loosely balanced regime can also generate realistic levels of variability. However, biologically there is no need for the network to generate all of its own variability, as all inputs to cortex are noisy (and there are other sources of noise such as stochasticity of cellular and synaptic mechanisms (Mainen and Sejnowski, 1995; O’Donnell and van Rossum, 2014; Schneidman et al., 1998) and input correlations (DeWeese and Zador, 2006; Stevens and Zador, 1998)). In at least one case (Sadagopan and Ferster, 2012), the noise derived from the cortical area’s input was shown to be large enough to potentially fully account for the noise seen in the cortical neurons. In the SSN network, the network will amplify input noise in the weakly coupled regime, and then decrease noise for increasingly strong inputs in the strongly coupled, loosely balanced

regime; the result is that, for higher input strengths, noise can be reduced to the level driven by the inputs (e.g., see Fig. 2D of Hennequin et al., 2018), consistent with the observations of Sadagopan and Ferster (2012).

In conclusion, we believe that at least sensory, and perhaps all of, cortex operates in a regime in which the inhibition and excitation neurons receive are loosely balanced. This along with the supralinear input/output function of individual neurons and simple assumptions on connectivity explains a large set of cortical response properties. A key outstanding question is the computational function or functions of this loosely balanced state and the response properties it creates (e.g., see Echeveste et al., 2019; G. Barello and Y. Ahmadian, in preparation).

## Appendix 1: Nomenclature for balanced solutions

There is no standard nomenclature for describing balanced solutions. Here we have used loose vs. tight balance to describe, given systematic cancellation, whether the remainder after cancellation is comparable to, or much smaller than, the factors that cancel.

Deneve and Machens (2016) used loose balance to mean that fast fluctuations of excitation and inhibition are uncorrelated, although they are balanced in the mean, as in the sparsely-connected network of van Vreeswijk and Sompolinsky (1998); and used tight balance to mean that fast fluctuations of excitation and inhibition are tightly correlated with a small temporal offset, as in the densely-connected network of Renart et al. (2010) and in the spiking networks of Deneve, Machens and colleagues in which recurrent connectivity has been optimized for efficient coding (Barrett et al., 2013; Boerlin et al., 2013; Bourdoukan et al., 2012). All of these networks are tightly balanced under our definition.

Hennequin et al. (2017) also defined balance to be tight if it occurs on fast timescales, and loose otherwise, but they implied that this is equivalent to our definition, that tight balance means the remainder is small compared to the factors that cancel, and loose balance means the remainder is comparable to the factors that cancel. The implied equivalence rests on the fact that tight balance under our definition produces very large (i.e.,  $O(\sqrt{K})$ ) negative eigenvalues (in linearization about the fixed point) which means very fast dynamics, approaching instantaneous population response as  $K$  and hence negative

eigenvalues go to infinity. We point out, however, that loose balance under our definition can produce negative eigenvalues large enough to produce quite fast dynamics, with effective time constants on the order of single neurons’ membrane time-constant, or even as small as a few milliseconds, depending on parameters.

An additional source of confusion is that there are two forms of fast fluctuations, with different behaviors. Fast fluctuations can be shared (correlated) across neurons, or they can be independent. The large negative eigenvalues in tightly balanced networks (in our definition) affect shared fluctuations corresponding to eigen-modes in which the activities of excitatory and inhibitory neurons fluctuate coherently. Thus, shared fluctuations are balanced on fast time scales. By contrast, spatial activity patterns in which neurons fluctuate independently are largely unaffected by those eigenvalues, and need not be balanced.

Fluctuations due to changes in population mean rates of the external input are shared, and so this form of fluctuation is followed on fast time scales in all balanced networks (at finite rates in loosely balanced networks, and approaching instantaneous following in tightly balanced networks). Fluctuations also arise from network and external neuronal spiking noise. In networks with sparse connectivity (van Vreeswijk and Sompolinsky, 1998), this yields independent fluctuations in different neurons, and thus independent fluctuations of excitation and inhibition on fast time scales (though their means are balanced). In networks with dense connectivity (Renart et al., 2010), these spiking fluctuations become shared fluctuations due to common inputs arising from the dense connectivity, and so in these networks excitation and inhibition are balanced on fast time scales. To summarize, all balanced networks will balance shared fluctuations, such as those due to changing external input rates, on fast time scales; but excitation and inhibition can nonetheless be unbalanced on fast time scales in sparse networks, due to independent fluctuations induced by spiking noise.

To conclude, we would argue for a future standardized terminology for dynamically-induced balancing of excitation and inhibition, in which “loose” vs. “tight” balance refer to our definition as to whether the remainder after cancellation is comparable to, or much smaller than, the factors that cancel. We suggest the use of “temporal” vs. “mean” balance to refer to whether or not excitation

and inhibition are balanced on fast time scales, which depends on whether there are substantial shared input fluctuations across neurons. “Finite” vs. “instantaneous” time scales of balancing can distinguish whether relaxation rates – the rates of balancing shared fluctuations – are moderately-sized vs. very large.

## Appendix 2: When do balanced solutions arise?

We consider a rate model in which the neuron’s input/output function is described by some function  $f(x)$ , which is zero for  $x \leq 0$  and monotonically increasing for  $x \geq 0$ . Then the network’s steady-state firing rate  $\mathbf{r}_{SS}$  for a steady input  $\mathbf{I}$  is

$$\mathbf{r}_{SS} = f(\mathbf{W}\mathbf{r}_{SS} + \mathbf{I}) \quad (14)$$

where  $f$  acts element by element on its argument, that is,  $f(\mathbf{u})$  is a vector whose  $i^{\text{th}}$  element is  $f(u_i)$  (the  $f$ ’s might differ for different neurons, which we neglect for simplicity). As before, we let  $\psi = \|\mathbf{W}\|$  and  $c = \|\mathbf{I}\|$ . We define the dimensionless and  $\mathcal{O}(1)$  matrix  $\mathbf{J} = \mathbf{W}/\psi$  and vector  $\mathbf{g} = \mathbf{I}/c$ , so that  $\mathbf{J}$  and  $\mathbf{g}$  represent the relative synaptic strengths and relative input strengths, respectively, while their overall magnitudes and dimensions are in  $\psi$  and  $c$ . Then, as in Ahmadian et al. (2013), we can define the dimensionless variable  $\mathbf{y} = \frac{\psi}{c}\mathbf{r}$ , and Eq. 14 can be rewritten

$$\mathbf{y}_{SS} = \frac{\psi}{c}f(c\mathbf{J}\mathbf{y}_{SS} + \mathbf{g}) \quad (15)$$

Note that this equation ensures that  $\mathbf{y}_{SS} \geq 0$ . Note also that, when  $f(x) = (x)_+^p$  ( $(x)_+ = x, x \geq 0; = 0$ , otherwise), then this equation can be rewritten  $\mathbf{y}_{SS} = \alpha(\mathbf{J}\mathbf{y}_{SS} + \mathbf{g})_+^p$  where  $\alpha = \psi c^{p-1}$ . This is the origin of the dimensionless constant  $\alpha$  mentioned in the main text.

If we define  $f^{-1}(0) = 0$ , then because  $f$  is monotonically increasing for non-negative arguments, it is invertible over that range, i.e.  $f^{-1}(x)$  is defined for  $x \geq 0$ . We can then rewrite Eq. 15 as

$$\frac{1}{c}f^{-1}\left(\frac{c}{\psi}\mathbf{y}_{SS}\right) = (\mathbf{J}\mathbf{y}_{SS} + \mathbf{g})_+ \quad (16)$$

If we assume

$$(\mathbf{J}\mathbf{y}_{SS} + \mathbf{g})_i \geq 0 \text{ for all } i, \quad (17)$$

that is, if  $(\mathbf{J}\mathbf{y}_{SS})_i \geq -(\mathbf{g})_i$  for all  $i$ , then we can replace the right side of Eq. 16 with  $\mathbf{J}\mathbf{y}_{SS} + \mathbf{g}$  (without the  $(\cdot)_+$ ). This condition, Eq. 17, is a condition on the solution  $\mathbf{y}_{SS}$ , which we must check is self-consistently met for any solution we derive under this assumption. Note also that,

from Eq. 15, the condition  $(\mathbf{J}\mathbf{y}_{SS} + \mathbf{g})_i > 0$  is met if and only if  $(\mathbf{y}_{SS})_i > 0$ , so if we find a solution  $\mathbf{y}_{SS}$  that has all positive elements, it will automatically satisfy Eq. 17. Given this assumption, a bit of further manipulation then yields

$$\mathbf{y}_{SS} = -\mathbf{J}^{-1}\mathbf{g} + \frac{1}{c}\mathbf{J}^{-1}f^{-1}\left(\frac{c}{\psi}\mathbf{y}_{SS}\right) \quad (18)$$

The first term,  $\mathbf{y}_{SS}^0 \equiv -\mathbf{J}^{-1}\mathbf{g}$ , is the balancing term, which cancels the external input  $\mathbf{g}$ , *i.e.*  $(\mathbf{J}\mathbf{y}_{SS}^0 + \mathbf{g}) = 0$ . If the second term becomes small relative to the first in some limit, then the tightly balanced solution,  $\mathbf{y}_{SS} \approx -\mathbf{J}^{-1}\mathbf{g}$  or equivalently  $\mathbf{r}_{SS} \approx -\mathbf{W}^{-1}\mathbf{l}$ , exists in that limit, while a loosely balanced solution (balance index  $\mathcal{O}(1)$ ) arises when the 2nd term is comparable in size to the first. (More careful analysis is needed to ensure that this solution is stable, and that there are not also other solutions.) Note that Eq. 18 gives an equation of the form Eq. 13 when we (1) Take  $f^{-1}(x) = (x)_+^{1/p}$  and (2) Multiply both sides of Eq. 18 by  $c/\psi$  to convert  $\mathbf{y}_{SS}$  to  $\mathbf{r}_{SS}$ .

Assuming all the elements of  $\mathbf{y}_{SS}^0 \equiv -\mathbf{J}^{-1}\mathbf{g}$  are  $> 0$ , a self-consistent solution in which the second term in Eq. 18 becomes small can be found in at least three cases:

- If  $c$  and  $\psi$  are scaled by the same factor, which becomes arbitrarily large, then there is a self-consistent solution in which  $\mathbf{y}_{SS}$  is converging to  $-\mathbf{J}^{-1}\mathbf{g}$ . Then the  $f^{-1}$  factor is not changing (except for the small changes due to the changes in  $\mathbf{y}_{SS}$  as it converges), but it is multiplied by the factor  $\frac{1}{c}$ , which becomes arbitrarily small; thus the second term becomes arbitrarily small, regardless of the particular structure of  $f$ . This is the case studied for the tightly balanced solution, where both  $c$  and  $\psi$  are taken proportional to  $\sqrt{K}$  with  $K$  very large. (Note that the mean field equations derived in (van Vreeswijk and Sompolinsky, 1998) differ from the generic steady-state rate equations, Eq. (14), in that they also involve the self-consistently calculated input fluctuation strengths,  $\sigma_A$ ; the scaling argument given here nevertheless holds in that case too.)
- Suppose  $c$  is scaled for fixed  $\psi$ , which is the biological case in which synaptic strengths are fixed and the strength of the external input is varied from small to large. Then if  $f^{-1}(x)$  grows more slowly than linearly in increasing  $x$ , then the  $\frac{1}{c}$  factor shrinks faster than the  $f^{-1}$  term grows, so

again there is a self-consistent solution in which  $\mathbf{y}_{SS}$  is converging to  $-\mathbf{J}^{-1}\mathbf{g}$  and the second term becomes arbitrarily small with increasing  $c$ . This is the case studied for the loosely balanced solution in the SSN, in which  $f(x)$  grows supralinearly with  $x$  and therefore  $f^{-1}(x)$  grows sublinearly with  $x$ .

- We again suppose  $c$  is scaled for fixed  $\psi$ , but now imagine that  $f^{-1}(x)$  grows faster than linearly in increasing  $x$ , *i.e.*  $f(x)$  is sublinear (for example,  $f(x) = (x)_+^p$  for  $0 < p < 1$ ). Then there is a self-consistent solution in which  $\mathbf{y} \rightarrow -\mathbf{J}^{-1}\mathbf{g}$  as  $c \rightarrow 0$ , with the second term in Eq. 18 going to zero as  $c \rightarrow 0$ . This case is the reverse of the SSN: the strongly coupled, balanced regime arises for  $c \rightarrow 0$ , while the weakly coupled, feedforward-driven regime arises for large  $c$ .

In sum, if the elements of  $-\mathbf{J}^{-1}\mathbf{g}$  are positive, then a self-consistent tightly-balanced solution arises for any  $f$  if  $c$  and  $\psi$  are scaled together by an increasing factor; for supralinear  $f$  if  $c$  is scaled by an increasing factor; or for sublinear  $f$  if  $c$  is scaled by a decreasing factor. In all of these cases, for moderate sizes of the scaled parameter(s) (*e.g.*, for the SSN, for  $\alpha = \mathcal{O}(1)$ ) such that the second term of Eq. 18 is comparable in size to the first, a loosely balanced solution should arise. Note that, since  $\mathbf{r}_{SS} = \frac{c}{\psi}\mathbf{y}_{SS}$ , then from Eq. 18 the net input after balancing should grow with increasing external input  $c$  as  $f^{-1}(c)$ ; this is sublinear in  $c$  for the SSN case of supralinear  $f$ .

If one considers a two-population model – a population of E cells and a population of I cells, with each population’s average rate represented by a single variable – then conditions on  $\mathbf{J}$  and  $\mathbf{g}$  can be defined such that the elements of  $-\mathbf{J}^{-1}\mathbf{g}$  are positive and the balanced fixed point is stable and is the only fixed point (Ahmadian et al., 2013; Kraynyukova and Tchumatchenko, 2018; van Vreeswijk and Sompolinsky, 1998). On the other hand, when the E element of  $-\mathbf{J}^{-1}\mathbf{g}$  is negative, Eq. 18 cannot serve as a basis for an asymptotic expansion with the leading term  $-\mathbf{J}^{-1}\mathbf{g}$ , and the tightly balanced state does not exist. (Given  $(-\mathbf{J}^{-1}\mathbf{g})_E < 0$ , if the tightly balanced state existed – meaning that the second term of Eq. 18 becomes much smaller than the first while Eq. 18 is applicable – then  $\mathbf{y}_{SS}$  must have crossed zero to become negative, but once that has happened we could no longer proceed past Eq. 16 and Eq. 18 would no longer be applicable, which is a contradiction; hence the tightly balanced state cannot exist.) However, the loosely balanced state can still arise

in this case in a broad parameter regime of  $\mathbf{J}$  and  $\mathbf{g}$ , and can be found as the fixed point of the iterative equation,  $\mathbf{y}_{SS}(t) = -\mathbf{J}^{-1}\mathbf{g} + \frac{1}{c}\mathbf{J}^{-1}f^{-1}\left(\frac{c}{\psi}\mathbf{y}_{SS}(t-1)\right)$ , given appropriate initial conditions  $\mathbf{y}_{SS}(0)$ ; see [Ahmadian et al. \(2013\)](#). In this case, with increasing  $c$ ,  $r_E$  grows, but then saturates and starts decreasing, and eventually is pushed down to 0. However, if we assume that the maximal external input (*i.e.* the maximal  $c$ ; for example, the maximal firing rate of the thalamic input to a primary sensory cortical area) can only drive  $r_E$  to saturation or slightly beyond, this represents a viable model of cortical systems ([Ahmadian et al., 2013](#); [Hennequin et al., 2018](#); [Rubin et al., 2015](#)).

A two-population model accurately describes the behavior of an unstructured model with many E and I neurons, *i.e.* with random connectivity and with neurons in each population receiving comparable stimulus inputs. In some cases this model also can form a good approximation to the behavior of a multi-neuron circuit with structured connectivity and stimulus selectivity ([Ahmadian et al., 2013](#)). More generally, though, in such a structured circuit with localized connectivity, for larger/stronger localized stimuli, some set of neurons (*e.g.*, neurons not selective for the stimulus) may eventually receive a net inhibition and become silent, meaning that the condition of Eq. 17 is not met and Eq. 18 does not apply. (However, if the connectivity is translation-invariant – the same at any position in the model – and the external input extends more narrowly than the network connections, then a balanced fixed point can still be attained, [Rosenbaum and Doiron, 2014](#).) Nonetheless, we find in simulations ([Ahmadian et al., 2013](#); [Hennequin et al., 2018](#); [Rubin et al., 2015](#)) that for reasonable stimulus input strengths, SSN behavior is reasonably described by the two-population model, in that (1) there is a transition with increasing input strength from a weakly coupled, feedforward-driven regime to a strongly-coupled, loosely balanced regime in which the input to excited neurons grows sublinearly as a function of the external input strength; and (2) if we define the  $\mathbf{W}$  and  $\mathbf{g}$  of the two-population model as describing the net input received by a cell in the larger, structured model – *e.g.*,  $W_{EE}$  represents the mean summed synaptic strength from excitatory cells to a single excitatory cell, and  $g_E$  represents the mean external input received by stimulus-selective excitatory cells – then reasonably good insight into the operating regime of the larger model can be obtained from the analysis of the two-population model presented

here and, in much more detail, in [Ahmadian et al. \(2013\)](#); [Kraynyukova and Tchumatchenko \(2018\)](#).

We believe the same overall analysis of a transition from a weakly coupled regime to a strongly coupled, loosely balanced regime will apply to multi-population models incorporating multiple subtypes of inhibitory cells (*e.g.* [Garcia Del Molino et al., 2017](#); [Kuchibhotla et al., 2017](#); [Litwin-Kumar et al., 2016](#)), but more detailed aspects of the analysis of the two-population model ([Ahmadian et al., 2013](#); [Kraynyukova and Tchumatchenko, 2018](#)) need to be generalized to that case.

## Acknowledgements

We thank Larry Abbott, Mario DiPoppa and Agostina Palmigiano for many helpful comments on the manuscript, David Hansel, Gianluigi Mongillo and Alfonso Renart for many useful discussions, and Randy Bruno for help with references. YA is supported by start-up funds from the University of Oregon. KDM is supported by NSF DBI-1707398, NIH U01NS108683, NIH R01EY029999, NIH U19NS107613, Simons Foundation award 543017, and the Gatsby Charitable Foundation.

## References

- Adesnik, H. (2017). Synaptic Mechanisms of Feature Coding in the Visual Cortex of Awake Mice. *Neuron*, 95:1147–1159.
- Ahmadian, Y., Rubin, D. B., and Miller, K. D. (2013). Analysis of the stabilized supralinear network. *Neural Computation*, 25:1994–2037.
- Amatrudo, J. M., Weaver, C. M., Crimins, J. L., Hof, P. R., Rosene, D. L., and Luebke, J. I. (2012). Influence of highly distinctive structural properties on the excitability of pyramidal neurons in monkey visual and prefrontal cortices. *J. Neurosci.*, 32(40):13644–13660.
- Amit, D. and Brunel, N. (1997). Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Comput. Neural Syst.*, 8:373–404.
- Anderson, J. S., Carandini, M., and Ferster, D. (2000). Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *J. Neurophysiol.*, 84:909–926.

- Anderson, J. S., Lampl, I., Gillespie, D. C., and Ferster, D. (2001). Membrane potential and conductance changes underlying length tuning of cells in cat primary visual cortex. *J. Neurosci.*, 21:2104–2112.
- Angelucci, A., Bijanzadeh, M., Nurminen, L., Federer, F., Merlin, S., and Bressloff, P. (2017). Circuits and mechanisms for surround modulation in visual cortex. *Ann. Rev. Neurosci.*, 40:425–451.
- Atallah, B. V. and Scanziani, M. (2009). Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. *Neuron*, 62:566–577.
- Barral, J. and Reyes, A. (2016). Synaptic scaling rule preserves excitatory-inhibitory balance and salient neuronal network dynamics. *Nat. Neurosci.*, 19:1690–1696.
- Barrett, D., Deneve, S., and Machens, C. (2013). Firing rate predictions in optimal balanced networks. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, pages 1538–1546. MIT Press.
- Barth, A. L. and Poulet, J. F. (2012). Experimental evidence for sparse firing in the neocortex. *Trends Neurosci.*, 35:345–355.
- Bhatia, A., Moza, S., and Bhalla, U. S. (2019). Precise excitation-inhibition balance controls gain and timing in the hippocampus. *Elife*, 8.
- Boerlin, M., Machens, C. K., and Deneve, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput. Biol.*, 9(11):e1003258.
- Bourdoukan, R., Barrett, D., Deneve, S., and Machens, C. (2012). Learning optimal spike-based representations. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, pages 2285–2293. MIT Press.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.*, 8:183–208.
- Carandini, M. and Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.*, 13:51–62.
- Cavanaugh, J. R., Bair, W., and Movshon, J. A. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J. Neurophysiol.*, 88:2530–2546.
- Chung, S. and Ferster, D. (1998). Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. *Neuron*, 20:1177–89.
- Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Sathianam, G., Sahani, M., and Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.*, 13:369–378.
- Cohen, M. R. and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.*, 14:811–819.
- Constantinople, C. M. and Bruno, R. M. (2013). Deep cortical layers are activated directly by thalamus. *Science*, 340:1591–1594.
- Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, R., Sader, E. N., Ko, H., Hofer, S. B., and Mrsic-Flogel, T. D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*, 518:399–403.
- Dehghani, N., Peyrache, A., Telenczuk, B., Le Van Quyen, M., Halgren, E., Cash, S. S., Hatsopoulos, N. G., and Destexhe, A. (2016). Dynamic Balance of Excitation and Inhibition in Human and Monkey Neocortex. *Sci Rep*, 6:23176.
- Deneve, S. and Machens, C. K. (2016). Efficient codes and balanced networks. *Nat. Neurosci.*, 19:375–382.
- DeWeese, M. and Zador, A. (2006). Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *J. Neurosci.*, 26:12206–12218.
- Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G. K., and Josic, K. (2016). The mechanics of state-dependent neural correlations. *Nat. Neurosci.*, 19:383–393.
- Echeveste, R., Aitchison, L., Hennequin, G., and Lengyel, M. (2019). Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. *bioRxiv*, doi: <http://dx.doi.org/10.1101/696088>.

- Ecker, A. S., Berens, P., Cotton, R. J., Subramanian, M., Denfield, G. H., Cadwell, C. R., Smirnakis, S. M., Bethge, M., and Tolias, A. S. (2014). State dependence of noise correlations in macaque primary visual cortex. *Neuron*, 82:235–248.
- Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., and Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science*, 327(5965):584–587.
- Elston, G. N. (2003). Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb. Cortex*, 13:1124–1138.
- Elston, G. N. and Fujita, I. (2014). Pyramidal cell development: postnatal spinogenesis, dendritic growth, axon growth, and electrophysiology. *Front Neuroanat*, 8:78.
- Elston, G. N. and Manger, P. (2014). Pyramidal cells in V1 of African rodents are bigger, more branched and more spiny than those in primates. *Front Neuroanat*, 8:4.
- Fares, T. and Stepanyants, A. (2009). Cooperative synapse formation in the neocortex. *Proc. Natl. Acad. Sci. U.S.A.*, 106(38):16463–16468.
- Feldmeyer, D., Egger, V., Lübke, J., and Sakmann, B. (1999). Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single “barrel” of developing rat somatosensory cortex. *J. Physiol.*, 521:169–190.
- Feldmeyer, D., Lübke, J., Silver, R. A., and Sakmann, B. (2002). Synaptic connections between layer 4 spiny neurone-layer 2/3 pyramidal cell pairs in juvenile rat barrel cortex: physiology and anatomy of interlaminar signalling within a cortical column. *J Physiol*, 538:803–822.
- Ferster, D., Chung, S., and Wheat, H. (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature*, 380:249–252.
- Finn, I. M., Priebe, N. J., and Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron*, 54:137–152.
- Galarreta, M. and Hestrin, S. (1998). Frequency-dependent synaptic depression and the balance of excitation and inhibition in the neocortex. *Nat. Neurosci.*, 1:587–594.
- Garcia Del Molino, L. C., Yang, G. R., Mejias, J. F., and Wang, X. J. (2017). Paradoxical response reversal of top-down modulation in cortical circuits with three interneuron types. *Elife*, 6.
- Graupner, M. and Reyes, A. D. (2013). Synaptic input correlations leading to membrane potential decorrelation of spontaneous activity in cortex. *J. Neurosci.*, 33:15075–15085.
- Haider, B., Duque, A., Hasenstaub, A. R., and McCormick, D. A. (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.*, 26:4535–4545.
- Haider, B., Hausser, M., and Carandini, M. (2013). Inhibition dominates sensory responses in the awake cortex. *Nature*, 493:97–100.
- Hansel, D. and van Vreeswijk, C. (2002). How noise contributes to contrast invariance of orientation tuning in cat visual cortex. *J. Neurosci.*, 22:5118–5128.
- Hansel, D. and van Vreeswijk, C. (2012). The mechanism of orientation selectivity in primary visual cortex without a functional map. *J. Neurosci.*, 32:4049–4064.
- Hennequin, G., Agnes, E. J., and Vogels, T. P. (2017). Inhibitory Plasticity: Balance, Control, and Codependence. *Annu. Rev. Neurosci.*, 40:557–579.
- Hennequin, G., Ahmadian, Y., Rubin, D. B., Lengyel, M., and Miller, K. D. (2018). The Dynamical Regime of Sensory Cortex: Stable Dynamics around a Single Stimulus-Tuned Attractor Account for Patterns of Noise Variability. *Neuron*, 98(4):846–860.
- Heuer, H. W. and Britten, K. H. (2002). Contrast dependence of response normalization in area MT of the rhesus macaque. *J. Neurophysiol.*, 88:3398–3408.
- Higley, M. and Contreras, D. (2006). Balanced excitation and inhibition determine spike timing during frequency adaptation. *J. Neurosci.*, 26:448–457.
- Histed, M. H. (2018). Feedforward Inhibition Allows Input Summation to Vary in Recurrent Cortical Networks. *eNeuro*, 5(1).
- Ichida, J. M., Schwabe, L., Bressloff, P. C., and Angelucci, A. (2007). Response facilitation from the “suppressive” receptive field surround of macaque V1 neurons. *J. Neurophysiol.*, 98:2168–2181.

- Kato, H. K., Asinof, S. K., and Isaacson, J. S. (2017). Network-Level Control of Frequency Tuning in Auditory Cortex. *Neuron*, 95(2):412–423.
- Ko, H., Hofer, S. B., Pichler, B., Buchanan, K. A., Sjostrom, P. J., and Mrsic-Flogel, T. D. (2011). Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473:87–91.
- Krainyukova, N. and Tchumatchenko, T. (2018). Stabilized supralinear network can give rise to bistable, oscillatory, and persistent activity. *Proc. Natl. Acad. Sci. U.S.A.*, 115(13):3464–3469.
- Kuchibhotla, K. V., Gill, J. V., Lindsay, G. W., Papadoyannis, E. S., Field, R. E., Sten, T. A., Miller, K. D., and Froemke, R. C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. *Nat. Neurosci.*, 20:62–71.
- Lankarany, M., Heiss, J. E., Lampl, I., and Toyozumi, T. (2016). Simultaneous Bayesian Estimation of Excitatory and Inhibitory Synaptic Conductances by Exploiting Multiple Recorded Trials. *Front Comput Neurosci*, 10:110.
- Li, L. Y., Li, Y. T., Zhou, M., Tao, H. W., and Zhang, L. I. (2013a). Intracortical multiplication of thalamocortical signals in mouse auditory cortex. *Nat. Neurosci.*, 16:1179–1181.
- Li, Y. T., Ibrahim, L. A., Liu, B. H., Zhang, L. I., and Tao, H. W. (2013b). Linear transformation of thalamocortical input by intracortical excitation. *Nat. Neurosci.*, 16:1324–1330.
- Lien, A. D. and Scanziani, M. (2013). Tuned thalamic excitation is amplified by visual cortical circuits. *Nat. Neurosci.*, 16:1315–1323.
- Litwin-Kumar, A., Chacron, M. J., and Doiron, B. (2012). The spatial structure of stimuli shapes the timescale of correlations in population spiking activity. *PLoS Comput. Biol.*, 8:e1002667.
- Litwin-Kumar, A., Rosenbaum, R., and Doiron, B. (2016). Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. *J. Neurophysiol.*, 115:1399–1409.
- Liu, L. D., Miller, K. D., and Pack, C. C. (2018). A Unifying Motif for Spatial and Directional Surround Suppression. *J. Neurosci.*, 38(4):989–999.
- Maimon, G. and Assad, J. A. (2009). Beyond Poisson: increased spike-time regularity across primate parietal cortex. *Neuron*, 62:426–440.
- Mainen, Z. F. and Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268:1503–1506.
- Marino, J., Schummers, J., Lyon, D. C., Schwabe, L., Beck, O., Wiesing, P., Obermayer, K., and Sur, M. (2005). Invariant computations in local cortical networks with balanced excitation and inhibition. *Nature Neurosci.*, 8:194–201.
- Markov, N. T., Misery, P., Falchier, A., Lamy, C., Vezoli, J., Quilodran, R., Gariel, M. A., Giroud, P., Ercsey-Ravasz, M., Pilaz, L. J., Huissoud, C., Barone, P., Dehay, C., Toroczkai, Z., Van Essen, D. C., Kennedy, H., and Knoblauch, K. (2011). Weight consistency specifies regularities of macaque cortical networks. *Cereb. Cortex*, 21(6):1254–1272.
- Markram, H., Lubke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol. (Lond.)*, 500 (Pt 2):409–440.
- Miller, K. D. (2016). Canonical computations of cerebral cortex. *Curr. Opin. Neurobiol.*, 37:75–84.
- Miller, K. D. and Troyer, T. W. (2002). Neural noise can explain expansive, power-law nonlinearities in neural response functions. *J. Neurophysiol.*, 87:653–659.
- Mongillo, G., Hansel, D., and van Vreeswijk, C. (2012). Bistability and spatiotemporal irregularity in neuronal networks with nonlinear synaptic transmission. *Phys. Rev. Lett.*, 108:158101.
- Nassi, J. J., Avery, M. C., Cetin, A. H., Roe, A. W., and Reynolds, J. H. (2015). Optogenetic Activation of Normalization in Alert Macaque Visual Cortex. *Neuron*, 86:1504–1517.
- Nienborg, H., Hasenstaub, A., Nauhaus, I., Taniguchi, H., Huang, Z. J., and Callaway, E. M. (2013). Contrast dependence and differential contributions from somatostatin- and parvalbumin-expressing neurons to spatial integration in mouse V1. *J. Neurosci.*, 33:11145–11154.

- O'Donnell, C. and van Rossum, M. C. (2014). Systematic analysis of the contributions of stochastic voltage gated channels to neuronal noise. *Front Comput Neurosci*, 8:105.
- Ohshiro, T., Angelaki, D. E., and DeAngelis, G. (2013). A normalization model of multisensory integration accounts for distinct forms of cross-modal and within-modal cue integration by cortical neurons. *Program No. 360.19. 2013 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2013. Online.*
- Okun, M. and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat. Neurosci.*, 11:535–537.
- Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., and Ferster, D. (2009). Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62:578–592.
- Pehlevan, C. and Sompolinsky, H. (2014). Selectivity and sparseness in randomly connected balanced networks. *PLoS ONE*, 9:e89992.
- Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., and Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*, 391:580–584.
- Poulet, J. F. and Petersen, C. C. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature*, 454:881–885.
- Priebe, N., Mechler, F., Carandini, M., and Ferster, D. (2004). The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nat. Neurosci.*, 7:1113–22.
- Priebe, N. J. and Ferster, D. (2008). Inhibition, spike threshold and stimulus selectivity in primary visual cortex. *Neuron*, 57:482–497.
- Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science*, 327:587–590.
- Reynolds, J. H. and Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu. Rev. Neurosci.*, 27:611–647.
- Rosenbaum, R. and Doiron, B. (2014). Balanced networks of spiking neurons with spatially dependent recurrent connections. *Physical Review X*, 4(2):021039.
- Rosenbaum, R., Smith, M. A., Kohn, A., Rubin, J. E., and Doiron, B. (2017). The spatial structure of correlated neuronal variability. *Nat. Neurosci.*, 20:107–114.
- Rubin, D. B., Van Hooser, S. D., and Miller, K. D. (2015). The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*, 85:402–417.
- Sadagopan, S. and Ferster, D. (2012). Feedforward origins of response variability underlying contrast invariant orientation tuning in cat visual cortex. *Neuron*, 74:911–923.
- Sadeh, S. and Rotter, S. (2015). Orientation selectivity in inhibition-dominated networks of spiking neurons: effect of single neuron properties and network dynamics. *PLoS Comput. Biol.*, 11:e1004045.
- Sanzeni, A., Akitake, B., Goldback, H., Leedy, C., Brunel, N., and Histed, M. (2019). Inhibition stabilization is a widespread property of cortical networks. *bioRxiv*, doi: <https://doi.org/10.1101/656710>.
- Sato, T. K., Hausser, M., and Carandini, M. (2014). Distal connectivity causes summation and division across mouse visual cortex. *Nat. Neurosci.*, 17:30–32.
- Sceniak, M., Ringach, D. L., Hawken, M., and Shapley, R. (1999). Contrast's effect on spatial summation by macaque v1 neurons. *Nature Neurosci.*, 2:733–739.
- Schneidman, E., Freedman, B., and Segev, I. (1998). Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput*, 10:1679–1703.
- Schoonover, C. E., Tapia, J. C., Schilling, V. C., Wimmer, V., Blazeski, R., Zhang, W., Mason, C. A., and Bruno, R. M. (2014). Comparative strength and dendritic organization of thalamocortical and corticocortical synapses onto excitatory layer 4 neurons. *J. Neurosci.*, 34:6746–6758.
- Schwabe, L., Ichida, J. M., Shushruth, S., Mangapathy, P., and Angelucci, A. (2010). Contrast-dependence of surround suppression in Macaque V1: experimental testing of a recurrent network model. *Neuroimage*, 52:777–792.
- Sengpiel, F., Blakemore, C., and Sen, A. (1997). Characteristics of surround inhibition in cat area 17. *Exp. Brain Res.*, 116:216–228.

- Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.*, 18:3870–3896.
- Shu, Y., Hasenstaub, A., and McCormick, D. A. (2003). Turning on and off recurrent balanced cortical activity. *Nature*, 423:288–293.
- Shushruth, S., Ichida, J. M., Levitt, J. B., and Angelucci, A. (2009). Comparison of spatial summation properties of neurons in macaque V1 and V2. *J. Neurophysiol.*, 102:2069–2083.
- Softky, W. and Koch, C. (1993). The highly irregular firing of cortical-cells is inconsistent with temporal integration of random EPSPS. *J. Neurosci.*, 13:334–350.
- Song, X. M. and Li, C. Y. (2008). Contrast-dependent and contrast-independent spatial summation of primary visual cortical neurons of the cat. *Cerebral Cortex*, 18:331–336.
- Stevens, C. and Zador, A. (1998). Input synchrony and the irregular firing of cortical neurons. *Nat. Neurosci.*, 1:210–217.
- Tan, A. Y., Chen, Y., Scholl, B., Seidemann, E., and Priebe, N. J. (2014). Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature*, 509:226–229.
- Troyer, T. W. and Miller, K. D. (1997). Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. *Neural Comput.*, 9:971–983.
- Tsodyks, M. V. and Sejnowski, T. (1995). Rapid state switching in balanced cortical network models. *Network*, 6:111–124.
- Tsui, J. M. and Pack, C. C. (2011). Contrast sensitivity of MT receptive field centers and surrounds. *J. Neurophysiol.*, 106:1888–1900.
- van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274:1724–1726.
- van Vreeswijk, C. and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Computation*, 10:1321–1371.
- Wang, S., Miller, K. D., and van Hooser, S. B. (2019). Combined visual and patterned optogenetic stimulation of ferret visual cortex reveals that cortical circuits respond to independent inputs in a sublinear manner. *Cosyne Abstracts 2019, Lisbon, Portugal*, (to appear):1–15.
- Wehr, M. and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, 426:442–446.
- Wilent, W. B. and Contreras, D. (2005). Stimulus-dependent changes in spike threshold enhance feature selectivity in rat barrel cortex neurons. *J. Neurosci.*, 25:2983–2991.
- Wu, G. K., Arbuckle, R., Liu, B. H., Tao, H. W., and Zhang, L. I. (2008). Lateral sharpening of cortical frequency tuning by approximately balanced inhibition. *Neuron*, 58:132–143.
- Wu, G. K., Li, P., Tao, H. W., and Zhang, L. I. (2006). Nonmonotonic synaptic excitation and imbalanced inhibition underlying cortical intensity tuning. *Neuron*, 52:705–715.
- Xue, M., Atallah, B. V., and Scanziani, M. (2014). Equalizing excitation-inhibition ratios across visual cortical neurons. *Nature*, 511:596–600.
- Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., O’Shea, D. J., Sohal, V. S., Goshen, I., Finkelshteyn, J., Paz, J. T., Stehfest, K., Fudim, R., Ramakrishnan, C., Huguenard, J. R., Hegemann, P., and Deisseroth, K. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, 477:171–178.
- Zhou, M., Liang, F., Xiong, X. R., Li, L., Li, H., Xiao, Z., Tao, H. W., and Zhang, L. I. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat. Neurosci.*, 17:841–850.