

RESEARCH ARTICLE

# Imbalanced amplification: A mechanism of amplification and suppression from local imbalance of excitation and inhibition in cortical circuits

Christopher Ebsch<sup>1</sup>, Robert Rosenbaum<sup>1,2\*</sup>

**1** Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, Indiana, United States of America, **2** Interdisciplinary Center for Network Science and Applications, University of Notre Dame, Notre Dame, Indiana, United States of America

\* [Robert.Rosenbaum@nd.edu](mailto:Robert.Rosenbaum@nd.edu)



**OPEN ACCESS**

**Citation:** Ebsch C, Rosenbaum R (2018) Imbalanced amplification: A mechanism of amplification and suppression from local imbalance of excitation and inhibition in cortical circuits. *PLoS Comput Biol* 14(3): e1006048. <https://doi.org/10.1371/journal.pcbi.1006048>

**Editor:** Lyle J. Graham, Université Paris Descartes, Centre National de la Recherche Scientifique, FRANCE

**Received:** October 10, 2017

**Accepted:** February 22, 2018

**Published:** March 15, 2018

**Copyright:** © 2018 Ebsch, Rosenbaum. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data files are available from the figshare database at <https://doi.org/10.6084/m9.figshare.5933365.v1>.

**Funding:** This work was supported by National Science Foundation grants DMS-1517828, DMS-1654268, and DBI-1707400. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Understanding the relationship between external stimuli and the spiking activity of cortical populations is a central problem in neuroscience. Dense recurrent connectivity in local cortical circuits can lead to counterintuitive response properties, raising the question of whether there are simple arithmetical rules for relating circuits' connectivity structure to their response properties. One such arithmetic is provided by the mean field theory of balanced networks, which is derived in a limit where excitatory and inhibitory synaptic currents precisely balance on average. However, balanced network theory is not applicable to some biologically relevant connectivity structures. We show that cortical circuits with such structure are susceptible to an amplification mechanism arising when excitatory-inhibitory balance is broken at the level of local subpopulations, but maintained at a global level. This amplification, which can be quantified by a linear correction to the classical mean field theory of balanced networks, explains several response properties observed in cortical recordings and provides fundamental insights into the relationship between connectivity structure and neural responses in cortical circuits.

## Author summary

Understanding how the brain represents and processes stimuli requires a quantitative understanding of how signals propagate through networks of neurons. Developing such an understanding is made difficult by the dense interconnectivity of neurons, especially in the cerebral cortex. One approach to quantifying neural processing in the cortex is derived from observations that excitatory (positive) and inhibitory (negative) interactions between neurons tend to balance each other in many brain areas. This balance is achieved under a class of computational models called “balanced networks.” However, previous approaches to the mathematical analysis of balanced network models is not possible under some biologically relevant connectivity structures. We show that, under these structures, balance

**Competing interests:** The authors have declared that no competing interests exist.

between excitation and inhibition is necessarily broken and the resulting imbalance causes some stimulus features to be amplified. This “imbalanced amplification” of stimuli can explain several observations from recordings in mouse somatosensory and visual cortical circuits and provides fundamental insights into the relationship between connectivity structure and neural responses in cortical circuits.

## Introduction

Information about a sensory stimulus is passed along a hierarchy of neural populations, from subcortical areas to the cerebral cortex where it propagates through multiple cortical areas and layers. Within each layer, lateral synaptic connectivity shapes the response to synaptic input from upstream layers and populations. In a similar manner, lateral connectivity shapes the response of cortical populations to artificial, optogenetic stimuli. The densely recurrent structure of local cortical circuits can lead to counter-intuitive response properties [1–5], making it difficult to predict or interpret a population’s response to natural or artificial stimuli. This raises the question of whether there are underlying arithmetical principles through which one can understand the relationship between a local circuit’s connectivity structure and its response properties.

In principle this relationship could be deduced from detailed computer simulations of the neurons and synapses that comprise the circuit. In practice, such detailed simulations can be computationally expensive, depend on a large number of unconstrained physiological parameters, and their complexity can make it difficult to pinpoint mechanisms underlying observed phenomena. In many cases, however, one is not interested in how the response of each neuron is related to the detailed connectivity between every pair of neurons. Relevant questions are often more macroscopic in nature, *e.g.* “How does increased excitation to population A affect the average firing rate of neurons in population B?” For such questions, it is sufficient to establish a relationship between macroscopic connectivity structure and macroscopic response properties.

One such approach is provided by the mean-field theory of balanced networks [6–10], which is derived in the limit of a large number of neurons and a resulting precise balance of strong excitation with strong inhibition. This notion of precise balance implies a simple relationship between the macroscopic structure of connectivity and firing rates, and balanced network models naturally produce the asynchronous, irregular spiking activity that is characteristic of cortical recordings [6, 7, 11, 12]. However, classical balanced network theory has some critical limitations. While cortical circuits do appear to balance excitation with inhibition, this balance is not always as precise and spike trains are not as asynchronous as the theory predicts [13–19]. Moreover, precise balance is mathematically impossible under some biologically relevant connectivity structures [8–10], implying that the classical theory of balanced networks is limited in its ability to model the complexity of real cortical circuits.

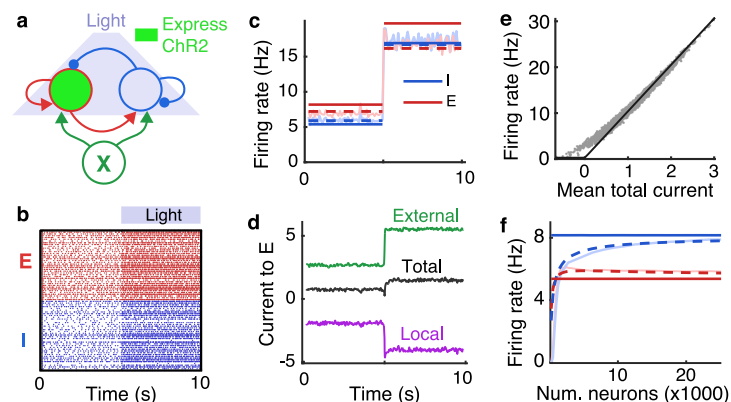
We show that cortical circuits with structure that is incompatible with balance are susceptible to an amplification mechanism arising when excitatory-inhibitory balance is broken at the level of local subpopulations, but maintained at a global level. This mechanism of “imbalanced amplification” can be quantified by a linear, finite-size correction to the classical mean field theory of balanced networks that accounts for imperfect balance and local imbalance. Through several examples, we show that imbalanced amplification explains several experimentally observed cortical responses to natural and artificial stimuli.

## Results

### The arithmetic of imprecise balance in cortical circuits

We begin by reviewing and demonstrating the classical mean-field theory of balanced networks and a linear correction to the large network limit that the theory depends on. A typical cortical neuron receives synaptic projections from thousands of neurons in other cortical layers, cortical areas or thalamus. These long range projections are largely excitatory and provide enough excitation for the postsynaptic neuron to spike at a much higher rate than the sparse spiking typically observed in cortex. The notion that excitation to cortical populations can be excessively strong has been posed in numerous studies and is typically resolved by accounting for local, lateral synaptic input that is net-inhibitory and partially cancels the strong, net-excitatory external synaptic input [2, 6, 20–23]. Balanced network theory takes this cancellation to its extreme by considering the limit of large external, feedforward synaptic input that is canceled by similarly large local, recurrent synaptic input. In this limit, a linear mean-field analysis determines population-averaged firing rates in terms of the macroscopic connectivity structure of the network [6, 7].

To demonstrate these notions, we first simulated a recurrent network of  $N_E = 4000$  excitatory (population  $E$ ) and  $N_I = 1000$  inhibitory spiking neurons (population  $I$ ) receiving synaptic connections from an “external” population ( $X$ ) of  $N_X = 4000$  excitatory neurons modeled as Poisson processes. Cortical circuits are often probed using optogenetic methods to stimulate or suppress targeted neuronal sub-populations [24, 25]. As a simple model of optogenetic stimulation of cortical pyramidal neurons, we added an extra inward current to all neurons in population  $E$  halfway through the simulation (Fig 1a). Neurons in the local population ( $E$  and  $I$ ) were modeled using the adaptive exponential integrate-and-fire (AdEx) model, which



**Fig 1. Imprecise balance under optogenetic stimulation.** **a**) Schematic. A population of recurrently connected excitatory (red) and inhibitory (blue) spiking neuron models receive synaptic input from an external population ( $X$ ; green) of Poisson-spiking neurons. Optogenetic stimulation of excitatory neurons was modeled by an extra inward current to the excitatory population at 5s. **b**) Spike rasters from 50 randomly selected excitatory (red) and inhibitory (blue) neurons from recurrent network. **c**) Average firing rate of excitatory (red) and inhibitory (blue) neurons in the recurrent network from simulations (light solid), from the balanced network approximation (Eq (3); solid dark) and from the corrected approximation (Eq (4); dashed). **d**) Mean synaptic currents to 200 randomly selected excitatory neurons in the recurrent network from external inputs ( $X$ ; green), from the local population ( $E + I$ ; purple) and the total synaptic current (black). Currents are measured in units of the neurons’ rheobase (rheobase/ $C_m = 10.5$  V/s). **e**) Mean firing rates plotted against mean input currents to all neurons in populations  $E$  and  $I$  (gray dots) and a rectified linear fit to their relationship (black line). **f**) Mean firing rates from identical simulations without stimulation except the total number of neurons,  $N$ , in the recurrent network was modulated while scaling synaptic weights and connection probabilities so that  $\epsilon \sim 1/\sqrt{N}$  (see Methods). Solid light curves are from simulations, solid dark from Eq (3), and dashed from Eq (4).

<https://doi.org/10.1371/journal.pcbi.1006048.g001>

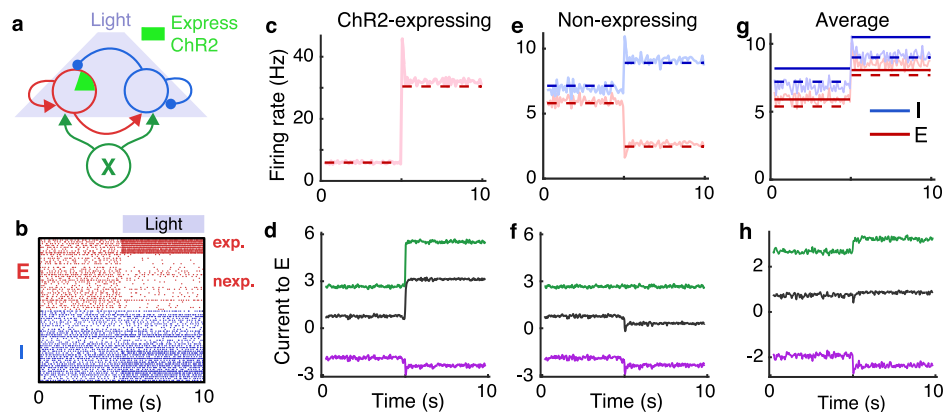
accurately captures the responses of real cortical neurons [26–28]. Connectivity was random with each neuron receiving 800 synaptic inputs on average and postsynaptic potential amplitudes between 0.19 and 1.0 mV in amplitude. The recurrent network produced asynchronous, irregular spiking (Fig 1b), similar to that observed in cortical recordings [11, 21, 29, 30]. Firing rates in populations *E* and *I* were similar in magnitude to those in population *X* and were increased by optogenetic stimulation (Fig 1c). As predicted by balanced network theory, local synaptic input (from *E* and *I* combined) was net-inhibitory and approximately canceled the external input from population *X* and artificial stimulation combined (Fig 1d).

**A review of the mean field theory of balanced networks.** To capture the notion that the net external synaptic input to neurons is strong, we define the small number,

$$\epsilon = \frac{1}{K_{EX}J_{EX}},$$

where  $K_{EX} = p_{EX} N_X$  is the average number of external synaptic projections received by each neuron in *E* from all neurons in *X*,  $p_{EX}$  is connection probability, and  $J_{EX}$  is the synaptic strength of each connection. Specifically,  $J_{EX}$  is the total postsynaptic current induced in a postsynaptic neuron in *E* by a single spike in a presynaptic neuron in *X*. Hence,  $1/\epsilon$  quantifies the synaptic current that would be induced in each neuron in *E* (on average) if every neuron in *X* spiked once simultaneously. For the simulations in Fig 1a–1e and in Fig 2,  $\epsilon = 0.0027$  mV. Using this convention, the mean synaptic input to each neuron in populations *E* and *I* from all sources can be written in vector form as

$$\mathbf{I} = \frac{1}{\epsilon} [W\mathbf{r} + \mathbf{X}]. \tag{1}$$



**Fig 2. Imbalanced amplification and suppression under partial optogenetic stimulation.** Same as Fig. 2a–d except the inward current was only provided to 20% of the excitatory neurons, modeling ChR2-expressing pyramidal cells. **a)** Schematic. A population of recurrently connected excitatory (red) and inhibitory (blue) spiking neuron models receive synaptic input from an external population (*X*; green) of Poisson-spiking neurons. Optogenetic stimulation of excitatory neurons was modeled by an extra inward current to 20% of the excitatory population at 5s. **b)** Spike rasters from 10 randomly selected ChR2-expressing and 40 non-expressing excitatory (red) neurons and 50 inhibitory (blue) neurons from the recurrent network. **c)** Average firing rate of ChR2-expressing excitatory neurons from simulations (light solid) and from the corrected approximation (Eq (4); dashed). **d)** Mean synaptic currents to 200 randomly selected ChR2-expressing excitatory neurons from external inputs (*X*; green), from the local population (*E* + *I*; purple) and the total synaptic current (black). **e)** Same as c, but for non-expressing excitatory neurons (red) and inhibitory neurons (blue). **f)** Same as d, but for non-expressing excitatory postsynaptic neurons. **g)** Same as c and d, but averaged over all excitatory neurons (expressing and non-expressing). Currents are measured in units of the neurons’ rheobase (rheobase/ $C_m = 10.5$  V/s). Firing rates predicted by Eq (3) are not shown in c and e because Eq (3) is not applicable to those cases.

<https://doi.org/10.1371/journal.pcbi.1006048.g002>

where  $\mathbf{I} = [I_E I_I]^T$  (superscript  $T$  denotes the transpose) is the vector of mean synaptic input to neurons in each population and similarly for their mean rates,  $\mathbf{r} = [r_E r_I]^T$ . The rescaled external synaptic input,  $\mathbf{X} = [X_E X_I]^T$ , is given by

$$\mathbf{X} = W_X r_X + \begin{bmatrix} s \\ 0 \end{bmatrix}.$$

where  $r_X$  is the average rate of neurons in population  $X$  and  $s/\epsilon$  is the strength of the inward current induced by optogenetic stimulation ( $s = 0$  when stimulation is off). The recurrent and feedforward mean-field connectivity matrices are given by

$$W = \begin{bmatrix} w_{EE} & w_{EI} \\ w_{IE} & w_{II} \end{bmatrix} \quad \text{and} \quad W_X = \begin{bmatrix} w_{EX} \\ w_{IX} \end{bmatrix}. \quad (2)$$

respectively where  $w_{ab} = K_{ab} J_{ab} / (K_{EX} J_{EX})$  quantifies the relative number,  $K_{ab} = p_{ab} N_b$ , and strength,  $J_{ab}$ , of synaptic connections from population  $b$  to  $a$ . To achieve moderate firing rates when  $\epsilon$  is small, local input,  $W\mathbf{r}$ , must be net-inhibitory and partially cancel the strong external excitation,  $\mathbf{X}$ , in Eq (1).

Balanced network theory [6, 7] takes this cancellation to its extreme by considering the limit of large number of neurons,  $N = N_E + N_I$ , while scaling connection strengths and probabilities in such a way that  $\epsilon \sim \mathcal{O}(1/\sqrt{N}) \rightarrow 0$ . Under this scaling, Eq (1) would seem to imply that mean synaptic currents diverge in the limit, but this divergence is avoided in balanced networks by a precise cancellation between external and recurrent synaptic input. To achieve this cancellation, firing rates must satisfy the mean-field balance equation,

$$W\mathbf{r} + \mathbf{X} = 0$$

in the large  $N$  limit, so that [6–10]

$$\mathbf{r} = -W^{-1}\mathbf{X}. \quad (3)$$

Hence, balanced network theory provides a closed form, linear expression for firing rates in the large network limit. Generally speaking, the firing rate of a neuron depends nonlinearly on the mean and variance of its input current [22, 31, 32]. Notably, however, the derivation of the fixed point in Eq (3) did not require us to specify the exact form of this dependence. Instead, Eq (3) represents the unique fixed point firing rates for which synaptic currents remain bounded as  $N \rightarrow \infty$ . More specifically, if Eq (3) is not satisfied as  $N \rightarrow \infty$  then  $\|\mathbf{I}\| \rightarrow \infty$  (where  $\|\cdot\|$  is the Euclidean norm). The existence of this fixed point does not guarantee that it is stable. Precise, general conditions on the accuracy of Eq (3) for spiking network models are not known and the investigation of such conditions is outside the scope of this study. However, the approximation tends to be accurate in the  $N \rightarrow \infty$  limit whenever all eigenvalues of  $W$  have negative real part, the solution in Eq (3) is strictly positive, and inhibitory synaptic kinetics are sufficiently fast [6–10, 32, 33]. Indeed, Eq (3) provides a reasonable, but imperfect approximation to firing rates in our spiking network simulation (Fig 1c, compare light and dark solid).

Balanced network theory has some critical limitations. Local cortical circuits are, of course, finite in size so the  $N \rightarrow \infty$  (equivalently  $\epsilon \rightarrow 0$ ) limit may not be justified. Moreover, excitation and inhibition in cortex may not be as perfectly balanced and spike trains not as asynchronous as predicted by balanced network theory [13–19, 34]. More importantly, under many biologically relevant connectivity structures, precise cancellation cannot be realized so Eq (3)

cannot even be applied [8–10]. We next review a simple, linear correction to Eq (4) that partially resolves these issues.

**A linear correction to precise balance.** A correction to Eq (3) can be obtained by considering  $\epsilon$  non-zero, but this requires making assumptions on the relationship between neurons' input statistics and firing rates. A simple approximation is obtained by assuming that population-averaged firing rates,  $\mathbf{r}$ , depend only on population-averaged mean inputs,  $\mathbf{I}$ , yielding the fixed points problem  $\mathbf{r} = f(\mathbf{I}) = f([\mathbf{W}\mathbf{r} + \mathbf{X}]/\epsilon)$  where  $f$  is the population-level f-I curve. When  $f$  is an increasing function over relevant ranges of  $\mathbf{I}$ , this fixed point equation can be re-written as

$$\mathbf{W}\mathbf{r} + \mathbf{X} = \epsilon f^{-1}(\mathbf{r}).$$

Hence, in strongly coupled networks ( $\epsilon$  small), the shape of f-I curves has a small effect on steady-state firing rates under such an approximation. Indeed, in the  $\epsilon \rightarrow 0$  limit, the f-I curve has no effect and firing rates are determined by Eq (3). This conclusion easily generalizes to the case where  $f$  also depends on the average temporal variance of neurons' inputs.

A simple case of this approximation is obtained by using a rectified linear approximation,  $\mathbf{r} = g[\mathbf{I}]_+$  where  $[\cdot]_+$  denotes the positive part. We fit such a function to the relationship between neurons' mean synaptic inputs and firing rates from our spiking network simulation (Fig 1e). Assuming that the average firing rates of all populations are positive, this rectified linear approximation produces a linear rate model [35] with mean firing rates given by solving  $\mathbf{W}\mathbf{r} + \mathbf{X} = \epsilon/g \mathbf{r}$  to obtain

$$\mathbf{r} = [\epsilon D - \mathbf{W}]^{-1} \mathbf{X} \tag{4}$$

where

$$D = \begin{bmatrix} 1/g & 0 \\ 0 & 1/g \end{bmatrix}.$$

The AdEx neuron model used in our simulations has a nonlinear f-I curve (Fig 1e; gray dots) and its firing rate depends on all statistics of its input, not just the mean [26, 36]. Nevertheless, the linear approximation in Eq (4) was accurate in predicting firing rates in our simulations (Fig 1c, dashed), outperforming the balanced network approximation from Eq (3). This can be explained by the fact that the balanced approximation in Eq (3) is already somewhat accurate and the linear approximation in Eq (4) corrects for some of the error introduced by imperfect balance, even though the true dependence of  $\mathbf{r}$  on  $\mathbf{I}$  is nonlinear.

To further investigate the relative accuracy of Eqs (3) and (4), we repeated the spiking network simulations from Fig 1a–1d while proportionally scaling the number of neurons ( $N_E$ ,  $N_I$ , and  $N_X$ ) in each population and scaling connection weights and probabilities in such a way that  $\epsilon \sim 1/\sqrt{N}$  (see Methods). As predicted by balanced network theory, excitatory and inhibitory firing rates increased toward the limit in Eq (3) (Fig 1f, compare light and dark solid lines). The linear correction in Eq (4) tracks this increase in firing rates and is more accurate than the approximation in Eq (3), particularly for smaller  $N$  (Fig 1f, dashed). It is worth noting that, in applying Eq (4) to obtain the dashed curve in Fig 1f, we fixed the value of  $g$  to the one obtained from the simulation in Fig 1a–1e. Hence, a single estimate of the gain yields an accurate approximation even under different parameter values.

The predictive power of Eq (4) in these examples is, of course, limited by the fact that it was only applied after estimating the gain of the neurons using firing rates obtained in simulations.



Moreover, highly nonlinear f-I curves could introduce additional error. However, the purpose of Eq (4) in this work is to provide a first-order approximation to and general understanding of firing rates in networks under which Eq (3) cannot be applied. For these purposes, Eq (4) is sufficient.

### Imbalanced amplification under partial optogenetic stimulation

We next show that a more realistic model of optogenetic stimulation breaks the classical balanced state, providing a demonstrative and experimentally relevant example of imbalanced amplification and suppression that explains phenomena observed in recordings from mouse somatosensory cortex.

**Firing rates are increased by stimulating fewer neurons.** The model of optogenetic stimulation considered in Fig 1 is somewhat inaccurate since optogenetic stimulation of excitatory neurons is often incomplete. For example, only a fraction of cortical pyramidal neurons express the channelrhodopsin 2 (ChR2) protein targeted in many optogenetic experiments [3, 24, 37, 38]. To more accurately model optogenetic stimulation, we modified the example above so the extra inward current was provided to a randomly chosen 20% of the excitatory neurons (Fig 2a), modeling ChR2-expressing pyramidal cells. This change produced surprising results. The ChR2-expressing neurons increased their firing rates by a larger amount than they did when all excitatory neurons received the current (Fig 2b and 2c; compare to Fig 1b and 1c). Hence, counterintuitively, stimulating fewer neurons actually *amplifies* the effects of stimulation on the targeted cells. In contrast, non-expressing excitatory neurons were *suppressed* during stimulation and inhibitory neurons increased their rates, but by a smaller amount than they did under complete stimulation (Fig 2e; compare to Fig 1c).

Similar effects were observed in experiments by Adesnik and Scanziani [3]. In that study, pyramidal neurons in layers (L) 2/3 of mouse somatosensory cortex (S1) were stimulated optogenetically, but only about 23% of the pyramidal neurons expressed ChR2. During stimulation, non-expressing L2/3 pyramidal neurons were suppressed and inhibitory synaptic currents increased, implying an increase in inhibitory neuron firing rates.

To understand these effects, we first extended the mean-field theory above to account for multiple subpopulations by defining

$$\mathbf{r} = \begin{bmatrix} r_{exp} \\ r_{nexp} \\ r_I \end{bmatrix}$$

to be the vector of average firing rates for the ChR2-expressing (*exp*), non-expressing (*nexp*) excitatory neurons and inhibitory (*I*) neurons. The vector of average input to the network is again given by Eq (1) where

$$\mathbf{X} = W_X r_X + \begin{bmatrix} s \\ 0 \\ 0 \end{bmatrix},$$

$$W_X = [w_{EX} \ w_{EX} \ w_{IX}]^T,$$

$$W = \begin{bmatrix} qw_{EE} & (1-q)w_{EE} & w_{EI} \\ qw_{EE} & (1-q)w_{EE} & w_{EI} \\ qw_{IE} & (1-q)w_{IE} & w_{II} \end{bmatrix}$$

and  $q = 0.2$  represents the proportion of neurons that express ChR2.

Importantly,  $W$  is singular (*i.e.*, not invertible), so classical balanced network theory fails for this example since Eq (3) cannot be evaluated. More specifically, it is impossible for  $\mathbf{I}$  in Eq (1) to remain finite as  $\epsilon \rightarrow 0$  since there is no vector,  $\mathbf{r}$ , such that  $W\mathbf{r} = -\mathbf{X}$ . Intuitively, this can be understood by noting that expressing and non-expressing excitatory neurons receive the same local input on average (Fig 2d and 2f, purple), since local connectivity is not specific to ChR2 expression, but they receive different external input during stimulation (Fig 2d and 2f, green). Therefore, local synaptic input cannot simultaneously cancel the external input to both sub-populations, so the precise cancellation required by classical balanced network theory cannot be achieved (Fig 2d and 2f, black). A similar mechanism has been used to explain a lack of cancellation between positive and negative correlations in balanced networks [12, 39].

**Amplification in the nullspace: A general analysis.** We now give a general analysis of network responses when  $W$  is singular. The example of partial optogenetic stimulation is then considered as a special case. If  $W$  is a singular matrix then only vectors,  $\mathbf{X}$ , that are in the column space of  $W$  admit solutions to  $W\mathbf{r} + \mathbf{X} = 0$ . The column space of  $W$  is defined as the linear space of all vectors,  $\mathbf{u}$ , such that  $\mathbf{u} = W\mathbf{r}$  for some  $\mathbf{r}$ . The column space of a matrix,  $W$ , is the orthogonal complement of the nullspace of  $W^T$ . We can therefore decompose

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{X}_1$$

where  $\mathbf{X}_0 = \text{proj}_{N(W^T)} \mathbf{X}$  is the projection of  $\mathbf{X}$  onto the nullspace of  $W^T$  and  $\mathbf{X}_1 = \text{proj}_{C(W)} \mathbf{X}$  is the projection onto the column space of  $W$ . Moreover, note that  $\text{proj}_{N(W^T)} W\mathbf{r} = 0$  since  $W\mathbf{r}$  is in the column space of  $W$ . Therefore, the projection of the total input onto the nullspace of  $W^T$  is

$$\text{proj}_{N(W^T)} \mathbf{I} = \text{proj}_{N(W^T)} \frac{1}{\epsilon} [W\mathbf{r} + \mathbf{X}] = \frac{1}{\epsilon} \mathbf{X}_0. \tag{5}$$

Hence, the projection of the total synaptic input onto the nullspace of  $W^T$  is  $\mathcal{O}(1/\epsilon)$  whenever  $\mathbf{X}$  has an  $\mathcal{O}(1)$  component in the nullspace of  $W^T$ . Note that, despite the  $1/\epsilon$  term in Eq (1), the total synaptic input,  $\mathbf{I}$ , is  $\mathcal{O}(1)$  when balance is realized due to cancellation (as in Fig 1d). Hence, the singularity of  $W$  introduces large,  $\mathcal{O}(1/\epsilon)$  synaptic currents where they would not occur if  $W$  was non-singular. In other words, external input in the nullspace of  $W^T$  produces strong synaptic currents in the network. Importantly, this conclusion does not rely on any assumptions about neurons' f-I curves or other properties. This result is a fundamental property of balanced networks or, more generally, networks receiving strong feedforward input.

To understand the implications of this result on firing rates in the network, however, we must specify an f-I curve. We again consider the linear rate approximation quantified by Eq (4). Importantly, unlike Eq (3) from classical balanced network theory, the approximation in Eq (4) is applicable to this example because it accounts for imperfect cancellation between local and external inputs. Specifically, the regularized matrix,  $\epsilon D - W$ , is invertible so Eq (4) can be evaluated even though Eq (3) cannot. The resulting firing rate solution from Eq (4)



agrees well with spiking network simulations (Fig 2c and 2e). Hence, Eq (4) provides an accurate approximation to firing rates in networks to which classical balanced network theory is not applicable at all.

Eq (4) also provides a concise mathematical quantification of firing rates when  $W$  is singular. Namely, if  $\mathbf{X}_0, \mathbf{X}_1 \sim \mathcal{O}(1)$  then firing rates can be expanded as

$$\mathbf{r} = \frac{1}{\epsilon} \mathbf{r}_0 + \mathbf{r}_1 \tag{6}$$

where  $\mathbf{r}_0$  is in the nullspace of  $W$  and  $\mathbf{r}_0, \mathbf{r}_1 \sim \mathcal{O}(1)$ . To derive this result, first note that Eq (4) can be rewritten as

$$W\mathbf{r} + \mathbf{X} = \epsilon D\mathbf{r}. \tag{7}$$

If  $\mathbf{X}$  has components in the nullspace of  $W^T$  then we can project both sides of this equation onto this nullspace to obtain

$$\text{proj}_{N(W^T)}\mathbf{X} = \epsilon \text{proj}_{N(W^T)}D\mathbf{r}.$$

where we again used the fact that  $\text{proj}_{N(W^T)}W\mathbf{r} = 0$  since  $W\mathbf{r}$  is in the column space of  $W$ . Since  $\mathbf{X}_0 = \text{proj}_{N(W^T)}\mathbf{X}$  and  $D$  are assumed  $\mathcal{O}(1)$ , this equation is only consistent when  $\mathbf{r} \sim \mathcal{O}(1/\epsilon)$ . We can therefore decompose  $\mathbf{r} = (1/\epsilon)\mathbf{r}_0 + \mathbf{r}_1$  where  $\mathbf{r}_0, \mathbf{r}_1 \sim \mathcal{O}(1)$ . We next show that  $\mathbf{r}_0$  is in the nullspace of  $W$ . From Eq (7), we have

$$W \left[ \frac{1}{\epsilon} \mathbf{r}_0 + \mathbf{r}_1 \right] + \mathbf{X} = \epsilon D \left[ \frac{1}{\epsilon} \mathbf{r}_0 + \mathbf{r}_1 \right].$$

Isolating the  $\mathcal{O}(1/\epsilon)$  terms gives  $W\mathbf{r}_0 = 0$  and therefore  $\mathbf{r}_0$  is in the nullspace of  $W$ . In summary, components of external input in the nullspace of  $W^T$  partially break balance to evoke amplified firing rates in the nullspace of  $W$ .

In the special case that  $W$  has a one-dimensional nullspace, a more precise characterization of  $\mathbf{r}_0$  is possible. Let  $\mathbf{v}_0$  be in the nullspace of  $W$  with  $\|\mathbf{v}_0\| = 1$ . Note that  $W^T$  also has a one-dimensional nullspace (since  $W$  is a square matrix). Let  $\mathbf{v}_2$  be in the nullspace of  $W^T$  with  $\|\mathbf{v}_2\| = 1$ . Since  $\mathbf{r}_0$  is in the nullspace of  $W$ , we can write  $\mathbf{r}_0 = a\mathbf{v}_0$  for some scalar,  $a$ . Now, dot product both sides of Eq (7) by  $\mathbf{v}_2$  to obtain

$$\begin{aligned} \mathbf{v}_2 \cdot \mathbf{X} &= \epsilon \mathbf{v}_2 \cdot D\mathbf{r} \\ &= \mathbf{v}_2 \cdot D[\mathbf{r}_0 + \epsilon \mathbf{r}_1] \end{aligned}$$

where we have used that  $\mathbf{v}_2 \cdot W\mathbf{r} = 0$  since  $\mathbf{v}_2$  is in the nullspace of  $W^T$ , which is orthogonal to  $W\mathbf{r}$  in the column space of  $W$ . Keeping only  $\mathcal{O}(1)$  terms and making the substitution  $\mathbf{r}_0 = a\mathbf{v}_0$ , we get

$$a = \frac{\mathbf{v}_2 \cdot \mathbf{X}}{\mathbf{v}_2 \cdot D\mathbf{v}_0}$$

so that

$$\mathbf{r}_0 = \frac{\mathbf{v}_2 \cdot \mathbf{X}}{\mathbf{v}_2 \cdot D\mathbf{v}_0} \mathbf{v}_0, \tag{8}$$

yielding a concise expression for the amplified component of firing rates when  $W$  has a one-dimensional nullspace. Note that  $\mathbf{v}_0$  is in the nullspace of  $W$ , so this result is consistent with the more general conclusions above.

**Amplification in the nullspace under partial optogenetic stimulation.** For the specific example of partial optogenetic stimulation considered in Fig 2, the nullspace of  $W^T$  is spanned by  $\mathbf{v}_2 = (1/\sqrt{2})[1 \ -1 \ 0]^T$  and the projection of  $\mathbf{X}$  onto the nullspace of  $W^T$  is  $\mathbf{X}_0 = [s/2 \ -s/2 \ 0]^T$ . The nullspace of  $W$  is spanned by  $\mathbf{v}_0 = (1/\sqrt{q^2 + (1 - q)^2})[1 - q \ -q \ 0]^T$ . We therefore have  $\mathbf{r} = (1/\epsilon)\mathbf{r}_0 + \mathbf{r}_1$  where Eq (8) gives

$$\mathbf{r}_0 = gs \begin{bmatrix} 1 - q \\ -q \\ 0 \end{bmatrix}.$$

Hence, ChR2-expressing neurons are amplified and non-expressing neurons are suppressed by optogenetic stimulation, as observed in simulations. A more precise result is given by expanding the full approximation from Eq (4) to obtain

$$\mathbf{r}_{\text{on}} = \mathbf{r}_{\text{off}} + \frac{gs}{\epsilon} \begin{bmatrix} 1 - q \\ -q \\ 0 \end{bmatrix} + \mathcal{O}(s) \begin{bmatrix} q \\ q \\ qc \end{bmatrix} + \mathcal{O}(\epsilon). \tag{9}$$

Here,  $\mathcal{O}(s)$  is a constant proportional to  $s$ ,  $c = |w_{IE}/w_{II}|$  and  $\mathbf{r}_{\text{off}}$  is the vector of firing rates in the balanced,  $\epsilon \rightarrow 0$ , limit when stimulation is off ( $s = 0$ ). Specifically,  $\mathbf{r}_{\text{off}}$  is the unique vector that satisfies  $W\mathbf{r}_{\text{off}} + W_{XrX} = 0$ , which is solvable even though  $W$  is singular because  $W_X r_X$  is in the column space of  $W$ , so balance can be maintained when  $s = 0$ .

The  $gs/\epsilon$  term in Eq (9) quantifies the amplification and suppression observed in simulations: Non-expressing neurons are suppressed by stimulation since  $-q < 0$  and the response of ChR2-expressing neurons is amplified since  $1 - q > 0$  and  $s/\epsilon$  is large. Stimulating fewer excitatory neurons ( $q$  smaller), increases the rate of the stimulated neurons, as indicated by the  $1 - q$  factor. The  $\mathcal{O}(s)$  term shows why inhibitory neurons increase their rates by a smaller amount. In summary, the optogenetically induced suppression observed experimentally by Adesnik and Scanziani [3] is a generic feature of balanced or strongly coupled networks under partial stimulation.

**Local imbalance with global balance explains intralaminar suppression and interlaminar facilitation.** Interestingly, despite the break of balance at the level of ChR2-expressing and non-expressing subpopulations, global balance is maintained in this example. This can be understood by repeating the mean-field analysis above without partitioning neurons into ChR2-expressing and non-expressing sub-populations, thereby quantifying the global average of firing rate of all excitatory neurons. In particular, the average synaptic input,  $\mathbf{I} = [I_E \ I_I]^T$ , to excitatory and inhibitory neurons is given by Eq (1) where  $W$  and  $W_X$  are as in Eq (2), and

$$\mathbf{X} = W_X r_X + \begin{bmatrix} sq \\ 0 \end{bmatrix}$$

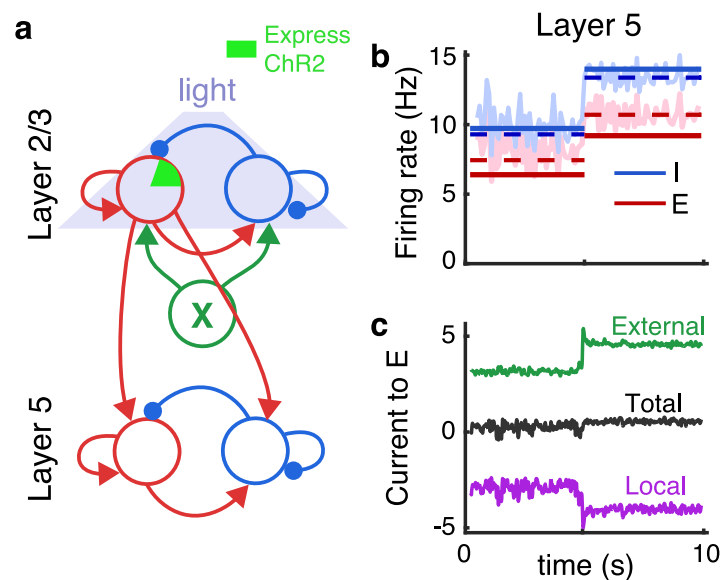
to account for the fact that only a proportion  $q$  of the excitatory neurons receive the inward current from optogenetic stimulation. In this case,  $W$  is non-singular so the balanced solution in Eq (3) is applicable. Indeed, the average firing rates of all excitatory neurons in our spiking network simulation is close to the prediction from Eq (3) and even closer to the prediction from Eq (4) (Fig 2g; compare to Fig 1c). The average feedforward input to all excitatory

neurons is canceled by net-inhibitory local input (Fig 2h; compare to Fig 1d). Hence, balance is maintained globally even though the network is imbalanced at the level of Chr2-expressing and non-expressing populations.

In the same study by Adesnik and Scanziani considered above [3], recordings were made in L5, which was not directly stimulated optogenetically, but receives synaptic input from L2/3. Interestingly, despite the fact that most excitatory neurons in L2/3 were suppressed during stimulation, firing rates in L5 increased.

To model these experiments, we interpreted the recurrent network from Fig 2 as a local neural population in L2/3, which sends synaptic projections to L5 (Fig 3a). We modeled a neural population in L5 identically to the L2/3 population, except its feedforward input came from excitatory neurons in the L2/3 network, instead of from Poisson-spiking neurons. As in experiments [3], L5 neurons increased their firing rates during stimulation (Fig 3b) and approximate balance was maintained (Fig 3c). This can be understood by noting that, in our model, L5 receives synaptic input sampled from all excitatory neurons in L2/3. Hence, the feedforward excitatory current to L5 neurons increases proportionally to the average excitatory firing rates in L2/3 during stimulation. As we showed above, this average rate increases (Fig 2e), despite the fact that most excitatory neurons in L2/3 are suppressed by stimulation. Hence, the combination of intralaminar suppression and interlaminar facilitation observed during optogenetic stimulation in experiments [3] results from the fact that the stimulated layer is locally imbalanced, but globally balanced during partial stimulation. These conclusions rely on an assumption that synaptic projections to L5 come from both Chr2-expressing and non-expressing excitatory neurons in L2/3.

**Imbalanced amplification of weak stimuli.** Sufficiently small  $\epsilon$  or large  $s$  would introduce negative rates in Eq (9), representing a regime in which non-expressing neurons cease spiking



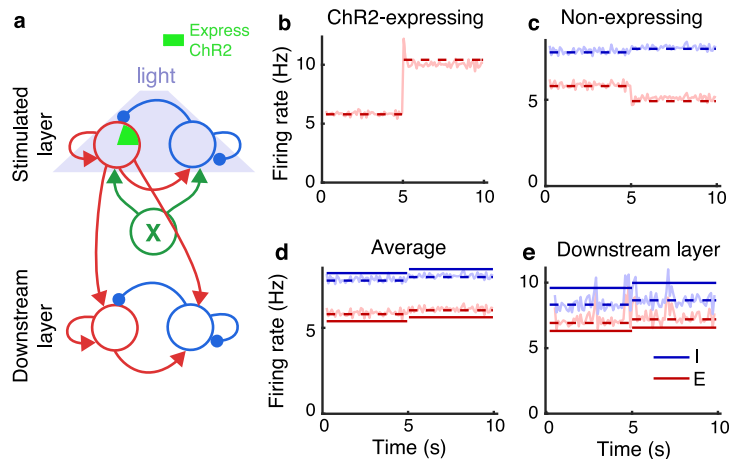
**Fig 3. Interlaminar facilitation despite intralaminar suppression under optogenetic stimulation.** a) Multi-layer network schematic. L2/3 was identical to the recurrent network in Fig 2 and provided external excitatory input to L5, which had the same internal structure as the L2/3 model. b) Average firing rates of excitatory (red) and inhibitory (blue) neurons in L5 from simulations (light solid) and from the corrected approximation (Eq (4); dashed). c) Mean synaptic currents to 200 randomly selected excitatory neurons in L5 from external inputs ( $X$ ; green), from the local population ( $E + I$ ; purple) and the total synaptic current (black). Currents are measured in units of the neurons' rheobase (rheobase/ $C_m = 10.5$  V/s).

<https://doi.org/10.1371/journal.pcbi.1006048.g003>

and the firing rate of ChR2-expressing neurons saturate at a high value. In this sense, firing rates do not truly have a  $\mathcal{O}(1/\epsilon)$  component for  $\epsilon$  very small. However, smaller values of  $\epsilon$  allow weak stimuli (small  $s$ ) to be strongly amplified. Strictly speaking, if one takes  $s \sim \mathcal{O}(\epsilon)$ , then under the linear approximation in Eq (4), partial optogenetic stimulation would have an  $\mathcal{O}(1)$  effect on the average firing rate of stimulated and unstimulated subpopulations, but an  $\mathcal{O}(\epsilon)$  effect on globally averaged firing rates. In practical terms, this means that, in strongly coupled networks ( $\epsilon$  small), partial optogenetic stimuli can have a moderate effect on the firing rates of stimulated neurons while having a negligible effect on the average firing rates of all excitatory neurons.

To demonstrate this idea, we repeated the simulations from Fig 2 in a network with four times as many neurons ( $N = 2 \times 10^4$ ) where synaptic weights and probabilities were scaled so that  $\epsilon \sim 1/\sqrt{N}$  (as in Fig 1f) and we reduced stimulus strength,  $s$ , as well. In this simulation, ChR2-expressing neurons' firing rates nearly doubled (Fig 4a) and non-expressing neurons were noticeably suppressed (Fig 4b). However, the change in the average firing rate of all excitatory neurons was nearly imperceptible (Fig 4c) and similarly for the firing rates of inhibitory neurons (Fig 4b and 4c). As a result, the firing rates in a downstream layer were unnoticeably modulated during stimulation (Fig 4d; compare to Fig 3). This effect could mask the effects of optogenetic stimulation in recordings.

**Imbalanced amplification with nearly singular connectivity matrices.** An apparent limitation of the results above is that they rely on the singularity of the connectivity matrix,  $W$ . Singularity is a fragile property of matrices that arises from structural symmetries. In the example above, singularity arises from our implicit assumption that local synaptic connectivity is independent of whether neurons express ChR2. Even a slight difference in connectivity to or from ChR2-expressing neurons would make  $W$  non-singular so that its nullspace would be



**Fig 4. Imbalanced amplification of weak stimuli.** Same simulations as Figs 2 and 3 except with  $N$  increased by a factor of four,  $\epsilon$  decreased by a factor of two, and a weaker stimulus. **a**) Schematic. A recurrent network (stimulated layer), which receives external input from Poisson-spiking neurons (green X) and from partial optogenetic stimulation, sends excitatory synaptic input to an identical network (downstream layer). **b**) Average firing rate of ChR2-expressing excitatory neurons in the stimulated layer from simulations (light solid) and from the corrected approximation (Eq (4); dashed). **c**) Same as b, but for non-expressing excitatory (red) and inhibitory (blue) neurons in the stimulated layer. **d**) Same as b, but averaged over all excitatory neurons in the stimulated layer. **e**) Same as d, but for the downstream layer. Mean firing rates from simulations of the stimulated layer changed from 5.8 Hz before stimulation to 10.0 Hz during stimulation for ChR2-expressing neurons, from 5.9 to 5.1 Hz for non-expressing excitatory neurons, from 5.9 to 6.1 Hz averaged over all excitatory neurons, and from 7.8 to 8.0 Hz for inhibitory neurons. Mean firing rates from simulations of the downstream layer changed from 7.2 Hz to 8.1 Hz for excitatory neurons and from 8.5 Hz to 9.5 Hz for inhibitory neurons.

<https://doi.org/10.1371/journal.pcbi.1006048.g004>

empty, rendering Eq (6) vacuous. We now show that Eq (6) and the surrounding analysis naturally extends to connectivity matrices that are approximately singular, with similar overall conclusions.

A matrix,  $W$ , is singular if it has  $\lambda = 0$  as an eigenvalue. A matrix can therefore be considered approximately singular if it has an eigenvalue with small magnitude. Specifically, let  $\lambda$  be an eigenvalue of  $W$  with  $|\lambda| \ll 1$ . Note that  $\lambda$  is also an eigenvalue of  $W^T$ . Now let  $\mathbf{v}$  be the associated eigenvector so that  $W^T \mathbf{v} = \lambda \mathbf{v}$  and assume that  $\|\mathbf{v}\| = 1$  without loss of generality. Take the projection of each term in Eq (7) onto the subspace spanned by  $\mathbf{v}$  to get

$$\text{proj}_{\mathbf{v}}[W\mathbf{r}] + \text{proj}_{\mathbf{v}}\mathbf{X} = \epsilon \text{proj}_{\mathbf{v}}[D\mathbf{r}].$$

Now note that  $\text{proj}_{\mathbf{v}}[W\mathbf{r}] = \lambda \text{proj}_{\mathbf{v}} \mathbf{r}$ . Hence,

$$\lambda \text{proj}_{\mathbf{v}} \mathbf{r} + \text{proj}_{\mathbf{v}}\mathbf{X} = \epsilon \text{proj}_{\mathbf{v}}[D\mathbf{r}].$$

If  $\text{proj}_{\mathbf{v}} \mathbf{X} \sim \mathcal{O}(1)$  and  $\text{proj}_{\mathbf{v}}[D\mathbf{r}] \sim \text{proj}_{\mathbf{v}} \mathbf{r}$  then this implies

$$(|\lambda| + \epsilon) \text{proj}_{\mathbf{v}} \mathbf{r} \sim \text{proj}_{\mathbf{v}} \mathbf{X}.$$

Hence,

$$\mathbf{r} = \frac{1}{\delta} \mathbf{r}_0 + \mathbf{r}_1$$

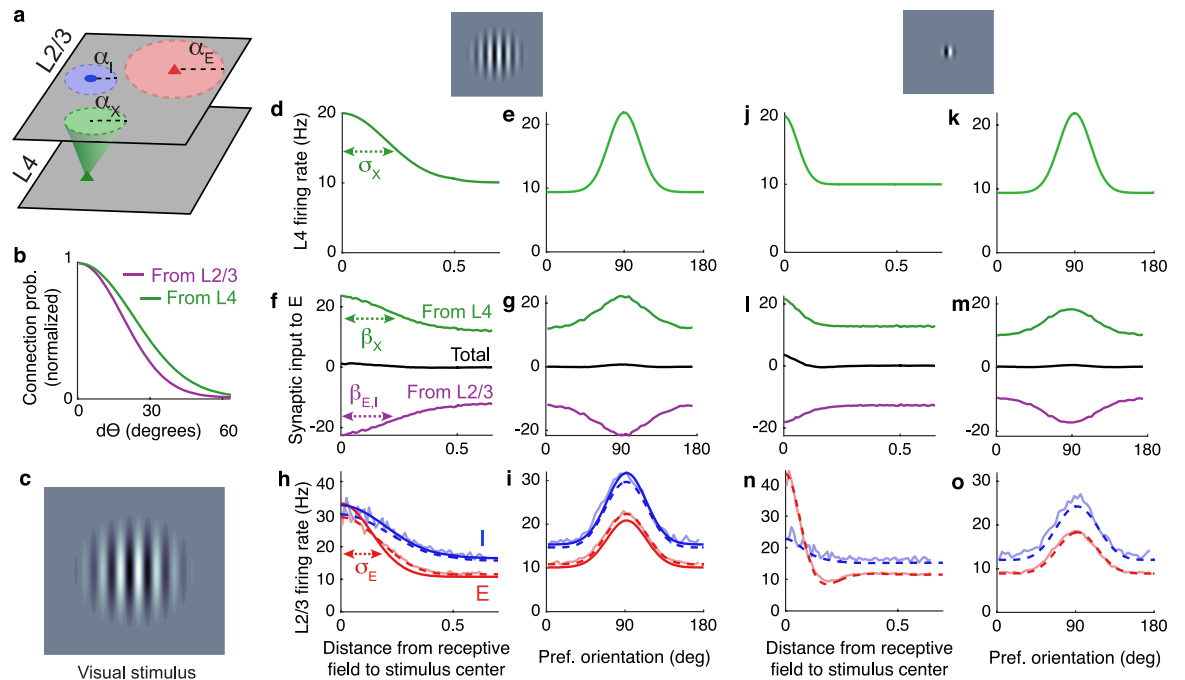
where  $\delta = |\lambda| + \epsilon$ . This generalizes Eq (6) to the case where  $W$  is only approximately singular. In summary, the mechanism of imbalanced amplification is a general property of strongly coupled networks with singular or nearly singular connection matrices.

We next show that networks with connection probabilities that depend on continuous quantities like distance or tuning preference necessarily have singular or nearly singular connectivity kernels and are therefore naturally susceptible to the amplification and suppression mechanisms described above.

### Imbalanced amplification and suppression in continuously indexed networks

So far we considered networks with discrete subpopulations. Connectivity in many cortical circuits depends on continuous quantities like distance in physical or tuning space. To understand how the amplification and suppression mechanisms discussed above extend to such connectivity structures, we next considered a model of a visual cortical circuit. We arranged  $2 \times 10^5$  AdEx model neurons (80% excitatory and 20% inhibitory) on a square domain, modeling a patch of L2/3 in mouse primary visual cortex (V1). Neurons received external input from a similarly arranged layer of  $1.6 \times 10^5$  Poisson-spiking neurons, modeling a parallel patch of L4 (Fig 5a). We additionally assigned a random orientation preference to each neuron, modeling the “salt-and-pepper” distribution of orientation preferences in mouse V1. Connectivity was probabilistic and, as in cortex [40–42], inter- and intralaminar connections were more numerous between nearby and similarly tuned neurons. Specifically, connection probability decayed like a Gaussian as a function of distance in physical and orientation space (Fig 5b), where distance in both spaces was measured using periodic boundaries, *i.e.* wrapped Gaussians were used in place of regular Gaussians.

**Amplification and suppression in simulations with spatially narrow stimuli.** An oriented stimulus localized in the animal’s visual field (Fig 5c) was modeled by imposing firing rate profiles in L4 that were peaked at the associated location in physical and tuning space, again with a Gaussian profile (Fig 5d and 5e). This produced external input to L2/3 that was



**Fig 5. Response properties of a continuously indexed network.** **a**) Network diagram. Poisson spiking neurons in L4 ( $X$ ) provide external synaptic input to  $2 \times 10^5$  recurrently connected excitatory and inhibitory AdEx model neurons ( $E$  and  $I$ ) in L2/3. The spatial width of synaptic projections from population  $a = X, E, I$  is given by  $\beta_a$ . **b**) Neurons are assigned random orientations and connection probability also depends on the difference,  $d\theta$ , between neurons' preferred orientation. **c**) An oriented stimulus in the animal's visual field. **d,e**) The location of the stimulus is modeled by firing rates in L4 that are peaked at the location of the stimulus in physical and orientation space. **f,g**) Synaptic current to neurons in population  $E$  from the external network (green), the local network (purple) and total (black) as a function distance from the receptive field center and as a function of neurons' preferred orientation. **h,i**) Firing rate profiles of excitatory (red) and inhibitory (blue) neurons in the local network from simulations (light curves), classical balanced network theory (solid, dark curves; from Eq (13)) and under the linear correction (dashed; from Eq (17)) in physical and orientation space. **j-o**) Same as (d-i) except for a smaller visual stimulus, modeled by a narrower spatial firing rate profile in L4. Firing rates from Eq (13) are not shown in panels n and o because balance cannot be realized and Eq (13) cannot be applied when external input is narrower than recurrent connectivity (see main text, S1 Text, and [8]).

<https://doi.org/10.1371/journal.pcbi.1006048.g005>

similarly peaked, but was nearly perfectly canceled by net-inhibitory lateral input (Fig 5f and 5g). Excitatory and inhibitory firing rate profiles in L2/3 were also peaked at the associated location in physical and tuning space (Fig 5h and 5i), demonstrating that neurons in L2/3 were appropriately tuned to the stimulus.

A smaller visual stimulus was modeled by shrinking the spatial profile of firing rates in L4 while leaving the orientation-dependence of L4 rates unchanged (Fig 5j and 5k). As above, synaptic inputs and firing rate profiles were appropriately peaked in physical and orientation tuning space (Fig 5l–5o). However, the smaller stimulus produced a surprising change to firing rates in L2/3. Despite the fact that L2/3 neurons at all locations received less excitation from L4 (Fig 5l), peak firing rates in L2/3 increased and a surround suppression dynamic emerged (Fig 5n). Hence, a more localized external input produced an amplification and suppression dynamic similar to the one observed in our model of optogenetic stimulation (compare to Fig 2). On the other hand, responses in orientation tuning space were mostly unchanged by the smaller stimulus size (Fig 5m and 5o).

**Mean-field theory of balance in two-dimensional spatial networks with orientation-tuning-specific connectivity.** The mean-field theory of balanced networks was previously extended to continuously indexed networks in one and two dimensions [8, 12, 43]. We now review a straightforward extension to two spatial dimensions and one orientation dimension.



Eq (1) generalizes naturally to

$$\mathbf{I} = \frac{1}{\epsilon} [\mathcal{W}\mathbf{r} + \mathbf{X}] \tag{10}$$

where  $\mathbf{r}(\mathbf{x}, \theta) = [r_E(\mathbf{x}, \theta) \ r_I(\mathbf{x}, \theta)]^T$  is the vector of mean firing rates of excitatory and inhibitory L2/3 neurons near spatial coordinates  $\mathbf{x} = (x, y)$  with preferred orientation near  $\theta$ , and similarly for the neurons' external input,  $\mathbf{X}(\mathbf{x}, \theta)$ , and total input,  $\mathbf{I}(\mathbf{x}, \theta)$ . The external input is given by  $\mathbf{X} = \mathcal{W}_X r_X$  where  $r_X(\mathbf{x}, \theta)$  is the profile of firing rates in L4. The connectivity kernels,  $\mathcal{W}$  and  $\mathcal{W}_X$ , are convolution integral operators defined by

$$\mathcal{W}\mathbf{r} = \begin{bmatrix} w_{EE} * r_E + w_{EI} * r_I \\ w_{IE} * r_E + w_{II} * r_I \end{bmatrix}$$

and

$$\mathcal{W}_X r_X = \begin{bmatrix} w_{EX} * r_X \\ w_{IX} * r_X \end{bmatrix}.$$

Here,  $w_{ab}(\mathbf{x}, \theta)$  is the mean-field connection strength between neurons separated by  $\mathbf{x}$  in physical space and  $\theta$  in orientation space (see [Methods](#)), and  $[w_{ab} * r_b](\mathbf{x}, \theta)$  denotes circular convolution with respect to  $\mathbf{x}$  and  $\theta$ , *i.e.*, convolution with periodic boundaries. These convolution operators implement low-pass filters in orientation and physical space, capturing the effects of synaptic divergence and tuning-specific connection probabilities. Similar filters describe feed-forward connectivity in artificial convolutional neural networks used for image recognition [44].

Taking  $\epsilon \rightarrow 0$  in Eq (10) shows that that firing rates must satisfy

$$\mathcal{W}\mathbf{r} + \mathbf{X} = 0. \tag{11}$$

This is an analogue to Eq (7) for spatial networks. From here, one may be tempted to invert the integral operator  $\mathcal{W}$  to obtain a spatial analogue of Eq (3). However, integral operators are never invertible [45]. Specifically, since Eq (11) is an integral equation of the first kind, there necessarily exist external input profiles,  $\mathbf{X}(\mathbf{x}, \theta)$ , for which Eq (11) does not admit a solution so that the classical balanced state cannot be realized [8]. This implies that there always exist inputs that prevent a continuously indexed network from maintaining excitatory-inhibitory balance. To better understand why this is the case, we follow previous work [8, 12, 43, 46, 47] in transitioning to the spatial Fourier domain to rewrite Eq (11) as

$$\widetilde{W} \widetilde{\mathbf{r}} + \widetilde{\mathbf{X}} = 0. \tag{12}$$

Here,  $\widetilde{\mathbf{r}}(\mathbf{n}, k) = [\widetilde{r}_E(\mathbf{n}, k) \ \widetilde{r}_I(\mathbf{n}, k)]^T$  is a Fourier coefficient of  $\mathbf{r}(\mathbf{x}, \theta)$  and similarly for  $\widetilde{\mathbf{X}}(\mathbf{n}, k) = \widetilde{W}_X(\mathbf{n}, k) \widetilde{r}_X(\mathbf{n}, k)$  where  $\mathbf{n} = (n_1, n_2)$  is the two-dimensional spatial Fourier mode and  $k$  is the Fourier mode in tuning space. Importantly, the convolution operators above become ordinary matrices in the Fourier domain. Specifically,

$$\widetilde{W} = \begin{bmatrix} \widetilde{w}_{EE} & \widetilde{w}_{EI} \\ \widetilde{w}_{IE} & \widetilde{w}_{II} \end{bmatrix}$$

and

$$\widetilde{W}_X = \begin{bmatrix} \widetilde{w}_{EX} \\ \widetilde{w}_{IX} \end{bmatrix}$$

where  $\widetilde{w}_{ab}(\mathbf{n}, k)$  is a Fourier coefficient of  $w_a(\mathbf{x}, \theta)$ . Note that going from Eq (11) to Eq (12) requires that  $\mathcal{W}$  is a convolution operator and that the boundaries of the network are treated periodically, *i.e.*, the convolutions are circular.

Solving Eq (12) gives an analogue to Eq (3) for spatial networks in the Fourier domain,

$$\widetilde{\mathbf{r}} = -\widetilde{W}^{-1}\widetilde{\mathbf{X}}. \tag{13}$$

This equation gives all Fourier coefficients,  $\widetilde{\mathbf{r}}(\mathbf{n}, k)$ . However, this solution is only viable when the inverse transform exists, *i.e.*, when the Fourier series of  $\widetilde{\mathbf{r}}(\mathbf{n}, k)$  converges, which in turn requires that  $\|\widetilde{\mathbf{X}}(\mathbf{n}, k)\|$  converges to zero faster than  $\|\widetilde{W}(\mathbf{n}, k)\|$  as  $\mathbf{n} \rightarrow 0$  and  $k \rightarrow 0$ . More specifically,  $\widetilde{\mathbf{r}}(\mathbf{n}, k)$  in Eq (13) must be square-summable. Hence, balance can only be realized when recurrent connectivity, quantified by  $\widetilde{W}(\mathbf{n}, k)$ , has more power at high spatial frequencies than external input,  $\widetilde{\mathbf{X}}(\mathbf{n}, k)$ . In other words, for balance to be realized, external input,  $\mathbf{X}(\mathbf{x}, \theta)$ , cannot have “sharper” spatial features than the recurrent connectivity kernels,  $w_{ab}(\mathbf{x}, \theta)$  for  $a, b = E, I$ .

The use of Fourier series to solve Eq (11) relies on the translation invariance of the integral operator,  $\mathcal{W}$ , but the existence of external input profiles for which Eq (11) does not admit a solution is a more general feature that is true for any integral operators with Hilbert-Schmidt kernels. Specifically, the convolutions,  $w_{ab} * r_b$ , that define  $\mathcal{W}$  and  $\mathcal{W}_X$  can be replaced by integrals of the form  $\int w_{ab}(\mathbf{u}, \mathbf{v})r_b(\mathbf{v})d\mathbf{v}$  and the result that there exist  $\mathbf{X}(\mathbf{u})$  for which Eq (11) does not admit a solution is still true whenever the kernels,  $w_{ab}(\mathbf{u}, \mathbf{v})$  are finite (no delta functions) and square-integrable [45].

**Balance and imbalance in networks with Gaussian-shaped connectivity kernels.** A more intuitive understanding of when and why balance is broken is provided by considering the Gaussian-shaped connectivity profiles used in our spiking network simulations. This explanation applies equally to the spatial profile of firing rates and connectivity in physical and orientation space, so we do not distinguish between the two in this discussion. Similar calculations were performed previously for spatial networks [8], so we only review the results here and discuss some of their implications here. Details of the calculations are provided in S1 Text.

When L4 firing rates and all connectivity kernels are Gaussian-shaped, all firing rates and synaptic input profiles are also Gaussian-shaped in the balanced state (see [8] and S1 Text), so the analysis can be performed solely in terms of the widths of each of these profiles.

Let  $\sigma_a$  be the width of the firing rate profile in population  $a$ ,  $\alpha_a$  the width of outgoing synaptic connections from the presynaptic neurons in population  $a$ , and  $\beta_a$  the width of the spatial profile of synaptic input from population  $a = X, E, I$  (Fig 5a, 5d, 5f and 5h). Since L4 is the feedforward population in this context, we use  $X$  subscripts for L4 and  $E, I$  for L2/3 neurons.

Synaptic divergence broadens the profile of synaptic currents so that

$$\beta_a^2 = \sigma_a^2 + \alpha_a^2 \tag{14}$$

for  $a = E, I, X$ . This is due to the Gaussian shape of firing rates and connectivity, as well as the convolution that describes the resulting mean-field synaptic inputs (the convolution of a Gaussian with a Gaussian is Gaussian). For balance to be maintained, feedforward synaptic input from L4 must be precisely canceled by lateral synaptic input in L2/3. This, in turn,

requires that

$$\beta_E = \beta_I = \beta_X.$$

Combined with Eq (14), this implies that balance requires the widths of firing rate profiles in L2/3 to satisfy [8]

$$\begin{aligned} \sigma_E^2 &= \beta_X^2 - \alpha_E^2 \\ \sigma_I^2 &= \beta_X^2 - \alpha_I^2. \end{aligned} \tag{15}$$

This approximation accurately predicted firing rate profiles in our first spiking network simulation (Fig 5h and 5i, solid, dark curves have widths given by Eq (15)). Hence, by Eq (15), the requirement of cancellation in balanced networks implies that recurrent connectivity sharpens neurons' tuning, both in physical and orientation space.

Interestingly, Eq (15) implies that the amount by which excitatory and inhibitory firing rate profiles are sharpened in balanced networks is determined by the width of their *outgoing* synaptic projections. Pyramidal neurons in L2/3 of mouse V1 preferentially target similarly tuned neurons in L2/3, but the tuning of these lateral connection probabilities is much broader than the tuning of pyramidal neurons' firing rates [41] ( $\alpha_E > \sigma_E$  in orientation space). This observation is consistent with Eq (15): Excitatory neuron tuning curves are sharpened precisely because their outgoing connections are broadly tuned. Hence, sharpening of excitatory neuron tuning curves in L2/3 is naturally achieved in balanced networks with lateral excitation, without requiring lateral inhibition. Following the same line of reasoning, the broader orientation tuning of inhibitory neurons [40] ( $\sigma_I$  larger) suggests that they project more locally in orientation tuning space than pyramidal neurons ( $\alpha_I < \alpha_E$  in orientation space).

Eq (15) also clarify when and why balanced network theory fails for continuously indexed networks. If external inputs are sharper than lateral connectivity ( $\beta_X < \alpha_E$  or  $\beta_X < \alpha_I$ ) in physical or orientation space, then Eq (15) do not yield solutions for  $\sigma_E$  or  $\sigma_I$ . In other words, balance requires that

$$\alpha_E < \beta_X \text{ and } \alpha_I < \beta_X$$

because Eq (11) does not admit a solution when these inequalities are broken (see [8] and S1 Text). Intuitively, when the inequalities above are broken, recurrent connectivity paints with too broad a brush to cancel the sharper feedforward input, so balance cannot be realized. As a result, balanced network theory cannot be applied to the example in Fig 5j–5o with a smaller visual stimulus.

**A linear correction to balance quantifies amplification and suppression in continuously indexed networks.** We next derive a linear correction to Eq (13) that accounts for imperfect cancellation and, in doing so, gives firing rate approximations where classical balanced network theory fails. Specifically, we generalize the derivation of Eq (4) to continuously indexed networks. Under this linear approximation, firing rate profiles are given by solving

$$\mathcal{W}\mathbf{r} + \mathbf{X} = \epsilon\mathbf{g}\mathbf{r}. \tag{16}$$

This is an integral equation of the second kind, which generically admits firing rate solutions,  $\mathbf{r}$ , even when Eq (11) does not [45]. We again transition to the Fourier domain so Eq (16) becomes

$$\tilde{\mathbf{r}} = [\epsilon D - \tilde{\mathcal{W}}]^{-1} \tilde{\mathbf{X}}. \tag{17}$$

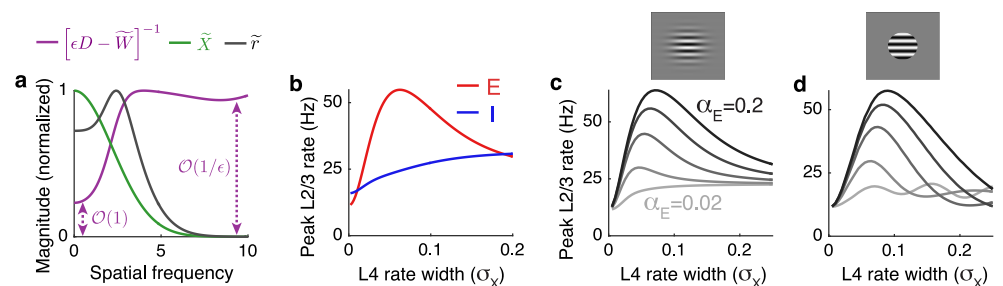
From Eq (17), firing rates,  $\mathbf{r}(\mathbf{x}, \theta)$ , can be computed numerically through an inverse transform

(the Fourier series over  $\mathbf{n}$  and  $k$ ), yielding an accurate approximation to firing rates from spiking network simulations even where classical balanced network theory fails (Fig 5n and 5o). Interestingly, the linear correction is slightly less accurate at predicting firing rates than the classical theory in Fig 5i. However, the main purpose of the corrected theory is not increased accuracy, but that it is applicable in situations where the classical theory is not.

The amplification and suppression caused by the smaller visual stimulus can be roughly explained by the balanced amplification mechanism discussed previously. Since  $\mathcal{W}$  is a low-pass filter, it approximately cancels high frequency components of firing rate profiles. Hence, high frequency components are in the approximate nullspace of the local connectivity operator,  $\mathcal{W}$ , and are therefore amplified by the network through the same mechanism discussed for discrete networks previously.

A more precise explanation is given by first averaging firing rates over orientation preference by setting  $k = 0$  in Eq (17) to give  $\tilde{\mathbf{r}}(\mathbf{n})$  that depends only on spatial frequency, and similarly for  $\tilde{\mathbf{X}}(\mathbf{n})$  and  $\tilde{\mathbf{W}}(\mathbf{n})$ . The convolution operator,  $\mathcal{W}$ , implements a low-pass filter, so  $\tilde{\mathbf{W}}(\mathbf{n})$  is  $\mathcal{O}(1)$  in magnitude at low spatial frequencies and converges to zero at higher frequencies (large  $\|\mathbf{n}\|$ ). The regularized inverse,  $[\epsilon D - \tilde{\mathbf{W}}(\mathbf{n})]^{-1}$ , is therefore  $\mathcal{O}(1)$  in magnitude at low frequencies and  $\mathcal{O}(1/\epsilon)$  at higher frequencies (Fig 6a, purple). When external input,  $\mathbf{X}(\mathbf{x})$ , has sharp features,  $\tilde{\mathbf{X}}(\mathbf{n})$  has power at higher spatial frequencies (Fig 6a, green), which are amplified by the  $\mathcal{O}(1/\epsilon)$  component of  $[\epsilon D - \tilde{\mathbf{W}}(\mathbf{n})]^{-1}$  while low frequencies remain  $\mathcal{O}(1)$ . The result is that the magnitude of  $\tilde{\mathbf{r}}(\mathbf{n})$  has a  $\mathcal{O}(1/\epsilon)$  peak at a non-zero spatial frequency (Fig 6a, black), introducing a high-amplitude, non-monotonic rate profile (as in Fig 5n; see [47] for a similar analysis). When  $\mathbf{X}(\mathbf{x})$  has spatially broad features,  $\tilde{\mathbf{X}}(\mathbf{n})$  has little power at high spatial frequencies so that this amplification dynamic is weak or absent (as in Fig 5h). An identical argument applies in orientation space. In summary, high-frequency components of external input profiles are transmitted more strongly than low-frequency components in strongly coupled networks, and the cutoff frequency is determined by the width ( $\alpha_E$  or  $\alpha_I$ ) of lateral synaptic projections.

It is worth noting that the average firing rates (over all orientations and spatial positions) are given by the zero Fourier coefficient  $\tilde{\mathbf{r}}(\mathbf{0}, 0)$ . When balance is broken by sharp external input features, the zero Fourier mode is not affected as long as mean firing rates,  $\mathbf{r}(\mathbf{x}, \theta)$ ,



**Fig 6. Spatial filtering of external input and the dependence of suppression on outgoing synaptic projection width.** **a** The magnitude of the spatial filter,  $[\epsilon D - \tilde{\mathbf{W}}(n)]^{-1}$ , imposed by recurrent connections (purple), the external input  $\tilde{\mathbf{X}}(n)$ , (green) and the resulting firing rate profile  $\tilde{\mathbf{r}}(n)$ , (black) as a function of the spatial frequency,  $\|\mathbf{n}\| = \sqrt{m^2 + n^2}$ , from the simulation in Fig 5j–5o. Magnitude is measured by the Frobenius norm for  $[\epsilon D - \tilde{\mathbf{W}}(n)]^{-1}$ . Curves normalized by their peaks. **b** Firing rates of excitatory (red) and inhibitory (blue) neurons with receptive fields at the center of a grating stimulus plotted as the width of the stimulus increases (represented by increasing  $\sigma_X$ ) using parameters from Fig 5j–5o. **c** Same as b, but the excitatory rate is plotted for different widths of the excitatory synaptic projection width,  $\alpha_E$ . **d** Same as c, but firing rates in L4 are shaped like a disc with radius  $\sigma_X$  instead of a Gaussian with width parameter  $\sigma_X$ .

<https://doi.org/10.1371/journal.pcbi.1006048.g006>

remain non-zero at all locations and orientations. Hence, sharp input features can break balance locally without breaking global, network-averaged, balance. This is analogous to the global balance obtained in the optogenetic example when local balance was broken at the level of subpopulations (Fig 2).

**Implications of imbalanced amplification on receptive field tuning.** We next considered a study by Adesnik et al. [48] in which drifting grating stimuli of varying sizes were presented to mice while recording from neurons in L2/3 of V1. In that study, pyramidal neurons' firing rates first increased then decreased as the stimulus size was increased. On the other hand, somatostatin-expressing (SOM) neuron's firing rates increased monotonically with stimulus size. Intracellular recordings combined with optogenetic stimulation in that study showed that SOM neurons project locally and pyramidal neurons form longer range projections.

To test our model against these findings, we applied Eq (17) to a network with local inhibition and longer-range excitation ( $\alpha_E > \alpha_I$ ) with increasing size of a visual stimulus (increasing  $\sigma_X$ ). Our results are consistent with recordings in Adesnik et al., 2012 [48]: Excitatory neuron firing rates changed non-monotonically with stimulus size, while inhibitory neuron firing rates monotonically increased (Fig 6b). The non-monotonic dependence of excitatory firing rates on stimulus size in Fig 6b is explained by the mechanism of imbalanced amplification. When  $\sigma_X$  is sufficiently small, balance is broken so imbalanced amplification introduces a large peak firing rate surrounded by suppression (as in Fig 5n). However, the total amount of external excitation introduced by the stimulus is proportional to the size of the stimulus, so a very small  $\sigma_X$  introduces very little excitation and peak firing rates are small. As  $\sigma_X$  increases, more excitation is recruited and the network is still imbalanced, which leads to increasingly large peak firing rates (as in Fig 5n). Once  $\sigma_X$  becomes large, balance begins to be restored and the peak excitatory firing rate decreases to moderate values (as in Fig 5h).

The degree to which excitatory neurons suppress depends on the spatial width,  $\alpha_E$ , of lateral excitatory projections (Fig 6c) and suppression of inhibitory neurons similarly depends on the spatial width,  $\alpha_I$ , of lateral inhibition (not pictured). Specifically, suppression occurs when lateral connectivity is broader than feedforward input ( $\alpha_E > \beta_X$  or  $\alpha_I > \beta_X$ ) because this is when the balanced solution in Eq (15) disappears. When a sub-population's lateral connectivity is more localized than feedforward connectivity from L4 ( $\alpha_E < \alpha_X$  as in the lightest gray curve in Fig 6c; or  $\alpha_I < \alpha_X$ ), that sub-population cannot exhibit suppression since feedforward input width ( $\beta_X^2 = \alpha_X^2 + \sigma_X^2$ ) is always larger than lateral connectivity, regardless of the stimulus size ( $\sigma_X$ ).

A similar line of reasoning explains why peak inhibitory neuron firing rates increase monotonically with stimulus size in Fig 6b. Inhibitory neurons in that example project locally ( $\alpha_I = \alpha_X$ ), so the inequality  $\alpha_I < \beta_X$  is always satisfied because  $\beta_X = \sqrt{\alpha_X^2 + \sigma_X^2} > \alpha_I$  whenever  $\alpha_I \leq \alpha_X$ . Whenever  $\alpha_I < \beta_X$ , inhibitory firing rates reflect their balanced state values which increase monotonically with the increase in total excitation induced by a larger stimulus.

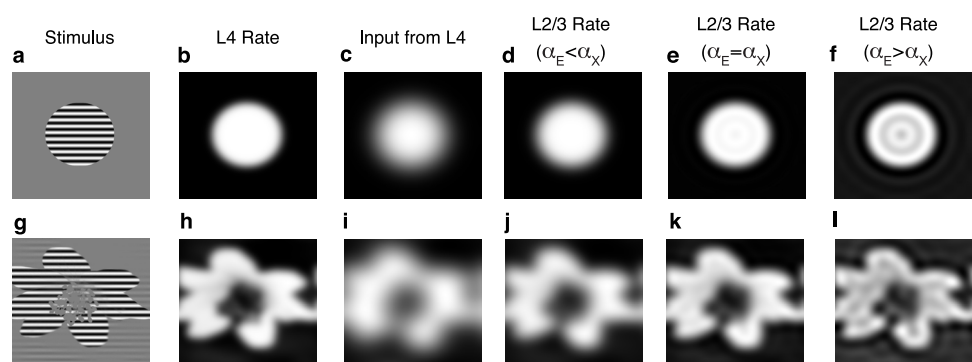
Unlike SOM neurons, parvalbumin-expressing (PV) neurons were found to exhibit suppression by Adesnik et al., 2012 [48]. Hence, our theory predicts that PV neurons project more broadly in space than SOM neurons. Indeed, PV interneurons in L2/3 are primarily basket cells whose axons project to larger lateral distances than other inhibitory neuron subtypes such as Martinotti cells that comprise most SOM neurons [49].

We observed a unimodal dependence of firing rate on stimulus size (Fig 6c, all curves have a single peak). However, Rubin, et al. [47] observed a multi-modal, oscillatory dependence of firing rate on stimulus size in recordings and in a computational model. In that study, the drifting grating stimuli were disc-shaped with a sharp cutoff of contrast at the edges of the

disc. Above, we considered a Gaussian-shaped contrast profile with soft edges (Fig 6c, inset). Repeating our calculations with a sharp-edged, disc-shaped stimulus (Fig 6d, inset) produced an oscillatory dependence of firing rate on stimulus size (Fig 6d), as observed by Rubin *et al.* This oscillation only arose when lateral synaptic projections were narrower than the stimulus size ( $\alpha_E$  small). The oscillation results from a Gibbs phenomenon: The sharp edge in the stimulus produces high-frequency power in  $\tilde{X}$ , which passes through the high-pass filter  $[\epsilon D - \tilde{W}]^{-1}$  when  $\alpha_E$  is small.

We next explored the functional consequences of these results on receptive field tuning. We first considered a disc-shaped grating stimulus (Fig 7a), producing a disc-shaped firing rate profile in L4 (Fig 7b). Synaptic divergence causes the profile of synaptic input from L4 to L2/3 to be “blurred” at the edges (Fig 7c), as quantified by the low-pass filter,  $\mathcal{W}_X$ . This illustrates a fundamental problem in receptive field tuning: Synaptic divergence from one layer to another implements a low-pass filter that blurs sharp features. This problem is resolved by our observation above that lateral, recurrent connectivity implements a high-pass filter. If the width of lateral, excitatory connections in L2/3 is similar to that of feedforward connections from L4, the high-pass filter implemented by the recurrent network cancels the low-pass filter implemented by feedforward connectivity, effectively implementing a deconvolution that can recover the sharpness of firing rate profiles in L4 (Fig 7d–7f). Hence, counterintuitively, *broader* lateral excitation actually *sharpen*s receptive field tuning. Broadening lateral connections further increases the sharpness of the firing rate profiles, but introduce oscillatory, Gibbs phenomena near sharp features (Fig 7f). These points are illustrated more clearly in an example with an asymmetrically shaped stimulus (Fig 7g–7l). Hence, the high-pass filter described above corrects the blurring caused by synaptic divergence between layers in V1.

In summary, imbalanced amplification and linear rate models provide a concise and parsimonious theoretical basis for understanding how suppression, amplification and tuning depends on the profile of neuron’s incoming and outgoing synaptic projections in physical and orientation tuning space.



**Fig 7. Imbalanced amplification and suppression reverse the blurring introduced by interlaminar synaptic divergence.** **a)** A disc-shaped grating stimulus gives rise to **b)** a disc-shaped firing rate profile,  $r_X(\mathbf{x})$ , in L4 with slightly blurred edges (achieved by convolving contrast from **a** with a Gaussian kernel). **c)** Input,  $X_E(\mathbf{x})$ , from L4 to excitatory neurons in L2/3 is blurred by synaptic divergence, which effectively applies a low-pass filter,  $\mathcal{W}_X$ , to the L4 rates. **d)** Excitatory firing rates in L2/3 are sharper than external input when lateral excitation is similar, but smaller, in width than interlaminar excitation ( $\alpha_E = 0.85\alpha_X$ ). **e)** Same as **c**, but lateral excitation is exactly as broad as interlaminar excitation ( $\alpha_E = \alpha_X$ ), which sharpens the edges further, making firing rates in L2/3 similar to those in L4. **f)** Same as **c**, but lateral excitation is broader than interlaminar excitation ( $\alpha_E = 1.1\alpha_X$ ), which sharpens the edges even further, but also introduced suppressed regions due to Gibbs phenomena. **g–l)** Same as **a–f**, but contrast was determined by the brightness of a photograph. Horizontal and vertical axes are neurons’ receptive fields.

<https://doi.org/10.1371/journal.pcbi.1006048.g007>



## Discussion

We described a theory of amplification in cortical circuits arising from a local imbalance that occurs when recurrent connectivity structure cannot cancel feedforward input. We showed that this imbalanced amplification is evoked by optogenetic stimuli in somatosensory cortex and sensory stimuli in visual cortex, since these stimuli cannot be canceled by the connectivity structure in those areas. Our theoretical analysis of imbalanced amplification explains several observations from cortical recordings in those areas.

Even though firing rates in balanced networks in the large  $N$  limit do not depend on neurons' f-I curves (see Eq (3)), quantifying firing rates under imbalanced amplification relies on a finite size correction that requires an assumption on how firing rates depend on neurons' input. For simplicity, we used an approximation that assumes populations' mean firing rates depend linearly on their average input currents, giving rise to Eqs (4) and (17). In reality, neurons' firing rates depend nonlinearly on their mean input currents, and also depend on higher moments of their input currents. However, the salient effects of imbalanced amplification are not sensitive to our assumption of linearity. For instance, Eq (5), which quantifies the strong synaptic currents evoked under imbalanced amplification, does not depend on any assumption about neurons' f-I curves. However, the precise value of the firing rates elicited by this strong input does depend on neurons' f-I curves. We found that the linear approximation to f-I curves in Eqs (4) and (17) performed well at approximating firing rates in our spiking network simulations and also explained several observations from cortical recordings. This may be partly explained due to the fact that our spiking network simulations used neuron models that exhibit spike frequency adaptation, which is known to linearize f-I curves [50, 51] and help networks maintain balance [10]. However, the linear approximation we used cannot explain some phenomena that rely on thresholding and other nonlinear transfer properties [47, 52]. The notion of imbalanced amplification extends naturally to models with nonlinear transfer functions and future work will consider the implications of nonlinearities.

Balanced networks are related to, but distinct from, inhibitory stabilized networks (ISNs) [2, 47, 53] and stabilized supralinear networks that can transition between ISN and non-ISN regimes [47]. The primary distinction is that ISNs are defined by moderately strong recurrent excitation (strong  $E \rightarrow E$ ) whereas balanced networks are defined by very strong external, feed-forward excitation (strong  $X \rightarrow E$ ) canceled by similarly strong net-inhibitory recurrent connectivity. Classical balanced networks are necessarily inhibitory stabilized at sufficiently large  $N$  (small  $\epsilon$ ) unless  $w_{EE} = 0$ . However, strongly coupled (approximately balanced) networks can be non-ISN at moderately large  $N$  (small  $\epsilon$ ) if  $w_{EE}$  is small. Cat V1 is believed to be inhibitory stabilized, which can be used to explain its surround suppression dynamic [2]. However, evidence from optogenetic and electrophysiological studies, suggests that mouse L2/3 V1 might not be inhibitory stabilized: Lateral connection probability is small between pyramidal neurons (small  $w_{EE}$ ) [49], stimulation of PV neurons does not produce the paradoxical effects that characterize ISNs [54], and modulating pyramidal neuron firing rates only weakly modulates excitatory synaptic currents in local pyramidal neurons [48, 54]. Nonetheless, pyramidal neurons and PV neurons in mouse V1 exhibit surround suppression [48], which we showed is explained by imbalanced amplification.

Despite the similarity in their names, the mechanism of imbalanced amplification studied here is fundamentally different from the mechanism of balanced amplification [55]. First, imbalanced amplification is related to steady-state firing rates, while balanced amplification is a dynamical phenomenon. Moreover, balanced amplification is intrinsic to the local, recurrent circuit: It produces large firing rate transients when local, recurrent inhibition is inefficient at canceling local, recurrent excitation. Imbalanced amplification, on the other hand, produces

large steady state firing rates when local, recurrent input is unable to effectively cancel feedforward, external excitation.

The analysis of our spatially extended network model relied on an assumption of periodic boundaries in space, which are not biologically realistic, but approximate networks with more realistic boundary conditions [8]. Without periodic boundary conditions, the integral eqs (10), (11) and (16) are equally valid, but the integrals are defined by regular convolutions in space instead of circular convolutions. As a result, the spatial Fourier modes do not de-couple, so Eqs (12), (13) and (17) are no longer valid, though they should still offer a good approximation when connectivity is much narrower than the the spatial domain [8]. In addition, anisotropic connectivity statistics, arising for example from tuning dependent connectivity in visual cortical circuits with coherent orientation maps [56], would prevent the integral operator in Eqs (10), (11) and (16) from being a convolution operator, and therefore preclude the use of Fourier series for the solution. Future work will consider the effects of non-periodic boundaries and non-convolutional connectivity kernels on spatially extended balanced networks.

We only considered neuron models with current-based synaptic input. In S2 Text, we show that our numerical results and analysis extend naturally to models with more realistic conductance-based synapses. The analysis makes use of an approximation that relates a conductance-based model to a current-based model with similar membrane potential statistics [57–59].

We focused on firing rates, but sensory coding also depends on variability and correlations in neurons’ spike trains. Our previous work derived the structure of correlated variability in heterogeneous and spatially extended balanced networks when connectivity structure prevents positive and negative correlations from cancelling, effectively providing an analogous theory of imbalanced amplification of correlated variability [12]. Combining those findings with the theory of steady-state firing rates presented here could yield a more complete theory of neural coding in cortical circuits and the effects of imbalanced amplification on coding.

## Materials and methods

We modeled recurrently connected networks with  $N$  neurons, composed of  $N_E = 0.8N$  excitatory and  $N_I = 0.2N$  inhibitory neurons. The recurrent network receives external input from a network of  $N_X$  neurons that drive the recurrent network. The membrane potential of neuron  $j$  from the excitatory ( $a = E$ ) or inhibitory ( $a = I$ ) population has Adaptive Exponential integrate-and-fire dynamics,

$$C_m \frac{dV_j^a}{dt} = -g_L(V - E_L) + g_L \Delta_T \exp[(V - V_T)/\Delta_T] + I_j^a(t) - w$$

$$\tau_w \frac{dw}{dt} = -w.$$

Whenever  $V_j^a(t) > V_{th}$ , a spike is recorded, the membrane potential is held for a refractory period  $\tau_{ref}$  then reset to a fixed value  $V_{re}$ , and  $w$  is incremented by  $B$ . Neuron model parameters for all simulations were  $\tau_m = C_m/g_L = 15\text{ms}$ ,  $E_L = -72\text{mV}$ ,  $V_T = -60\text{mV}$ ,  $V_{th} = -15\text{mV}$ ,  $\Delta_T = 1.5\text{mV}$ ,  $V_{re} = -72\text{mV}$ ,  $\tau_{ref} = 1\text{ms}$ ,  $\tau_w = 150\text{ms}$  and  $B = C_m 0.267 \text{ mV/ms}$ . We write all currents in terms of  $C_m$ , so  $C_m$  can be any constant. Membrane potentials were also bounded below by  $V_{lb} = -100\text{mV}$ . Synaptic input currents were defined by

$$I_j^a(t) = [X_j^a(t) + R_j^a(t)]C_m \tag{18}$$

where  $X_j^a(t)$  is the feedforward input and  $R_j^a(t)$  the recurrent input to neuron  $j$  in population

$a = E, I$ . The recurrent input was defined by

$$R_j^a(t) = \sum_{b=E,I} \sum_{k=1}^{N_b} J_{jk}^{ab} \sum_n \eta_b(t - t_n^{b,k})$$

where  $t_n^{b,k}$  is the  $n$ th spike time of neuron  $k$  in population  $b = E, I$ . The external input to the recurrent network is defined similarly by

$$X_j^a(t) = \sum_{k=1}^{N_X} J_{jk}^{aX} \sum_n \eta_X(t - t_n^{X,k}). \tag{19}$$

where  $t_n^{X,k}$  is the  $n$ th spike time of neuron  $k = 1, \dots, N_X$  in population  $X$ . Each coefficient,  $J_{jk}^{ab}$ , represents the synaptic weight from presynaptic neuron  $k$  in population  $b$  to postsynaptic neuron  $j$  in population  $a$ . For all simulations, we modeled synaptic kinetics using  $\eta_b(t) = \exp(-t/\tau_b)/\tau_b$  for  $t > 0$  where  $\tau_E = 8\text{ms}$ ,  $\tau_I = 4\text{ms}$ , and  $\tau_X = 10\text{ms}$ . Note that the integral of  $\eta_b(t)$  over time is equal to 1 for all three kernels, so the choice of time constant,  $\tau_b$ , does not effect time-averaged synaptic currents. We used  $\tau_I < \tau_E < \tau_X$  to prevent excessive synchronous events that break the balanced state. While inhibition may be faster than excitation in many cortical circuits, excitatory neurons are more likely to contact distal dendrites and inhibitory neurons are more likely to contact the soma [60, 61], which could make inhibition functionally faster than excitation. In any case, using fast inhibition is common practice in spiking network simulations with strong or dense connectivity [11, 12, 47, 62, 63] and a complete resolution of this issue is outside the scope of this study.

In Figs 1, 2 and 3 an extra term,  $S = 2 \text{ mV/ms}$ , was added to  $X_j^E(t)$  for stimulated neurons during the second half of the simulation to model optogenetic stimulation. We used  $N_E = 4000$ ,  $N_I = 1000$  and  $N_X = 4000$  (so  $N = 5000$ ) except for Fig 1f where all  $N_b$  values were scaled. Connections were drawn randomly with connection probabilities  $p_{EE} = p_{IE} = p_{IX} = 0.1$ ,  $p_{EI} = p_{II} = p_{EX} = 0.2$ . Specifically, for each neuron in presynaptic population  $b = E, I, X$ , we sampled  $p_{ab} N_a$  postsynaptic targets from population  $a = E, I$  randomly and uniformly with replacement. Since outgoing connections were sampled with replacement, some neurons connected multiple times to other neurons. Synaptic weights were then defined by

$$J_{jk}^{ab} = (\# \text{ of contacts}) \times J_{ab}$$

where  $J_{EE} = 0.4\text{mV}$ ,  $J_{IE} = 0.83 \text{ mV}$ ,  $J_{II} = J_{EI} = -1.67 \text{ mV}$ ,  $J_{EX} = J_{IX} = 0.47 \text{ mV}$ . This gives postsynaptic potential amplitudes between 0.19 and 1.0 mV. For Figs 1f and 4, the values of  $J_{ab}$  and the values of  $p_{ab}$  were each multiplied by  $(5000/N)^{1/4}$  so that they were unchanged at  $N = 5000$  and so that  $\epsilon \sim 1/\sqrt{N}$ . This is slightly different from the more common practice of fixing small connection probabilities and scaling  $J_{ab}$  like  $1/\sqrt{N}$ . We instead fixed a relatively dense connectivity at  $N = 5000$  and the network became increasingly sparse and weakly connected at increased  $N$ . Both approaches have the same mean-field (since the mean-field only depends on the product of  $p_{ab}$  and  $J_{ab}$ ), but our approach prevents excessively small synaptic weights at large  $N$  and prevents dense connectivity at large  $N$ , which is computationally expensive and susceptible to oscillatory and synchronous spiking.

Spike times in the external population were modeled as independent Poisson processes with  $r_X = 5 \text{ Hz}$ . In Fig 3, external input to the L5 population was created using the spike times of excitatory neurons from the simulations in Fig 2. Simulations for Fig 4 were identical to those in Figs 2 and 3 except there were  $N = 2 \times 10^4$  neurons in the L2/3 model, synaptic weights to neurons in that population were multiplied by  $1/\sqrt{2}$ , and connections probabilities were

also multiplied by  $1/\sqrt{2}$ . Hence, in relation to Fig 2,  $N$  was increased by a factor of four and  $\epsilon$  was halved.

Simulations for Fig 5 used algorithms adapted from previous work [12]. The recurrent network (L2/3) contained  $N = 2 \times 10^5$  AdEx model neurons,  $N_E = 1.6 \times 10^5$  of which were excitatory and  $N_I = 4 \times 10^4$  inhibitory. Excitatory and inhibitory neurons in L2/3 were arranged on a uniform grid covering the unit square  $[0, 1] \times [0, 1]$  (arbitrary spatial units). The external population (L4) contained  $N_X = 1.6 \times 10^5$  neurons arranged on an identical, parallel square. Each neuron in each population was assigned a preferred orientation chosen randomly and uniformly from 0 to  $180^\circ$ . Connections were chosen randomly as above, but connection probabilities depended on the neurons' distances in physical and orientation tuning space. Specifically, the connection probability from a neuron in population  $b = E, I, X$  at coordinates  $\mathbf{x} = (x_1, x_2)$  to a neuron in population  $a = E, I$  at coordinates  $\mathbf{y} = (y_1, y_2)$  was

$$p_{ab}(\mathbf{x} - \mathbf{y}, d\theta) = \bar{p}_{ab} G(\mathbf{x} - \mathbf{y}; \alpha_b) g(d\theta/180^\circ; \alpha_{b,\theta})$$

where  $d\theta$  is the difference between neurons' preferred orientation,

$$g(u; \alpha) = \frac{1}{\sqrt{2\pi\alpha}} \sum_{k=-\infty}^{\infty} e^{-u^2/(2\alpha^2)}$$

is a one-dimensional wrapped Gaussian and  $G(\mathbf{u}; \alpha) = g(u_1; \alpha)g(u_2; \alpha)$  is a two dimensional wrapped Gaussian. The connection probability averaged over all distances is  $\bar{p}_{ab}$ , which were chosen to be the same as in previous figures,  $\bar{p}_{EE} = \bar{p}_{IE} = \bar{p}_{IX} = 0.1$  and  $\bar{p}_{EI} = \bar{p}_{II} = \bar{p}_{EX} = 0.2$ . As above, outgoing connections were chosen with replacement, so some neurons made multiple contacts onto other neurons. Connection widths in physical space were  $\alpha_E = 0.15$  and  $\alpha_I = \alpha_X = 0.04$  (as measured on the unit square). Connection widths in orientation space were  $\alpha_{E,\theta} = \alpha_{E,\theta} = 0.1$  and  $\alpha_{X,\theta} = 0.125$  (corresponding to widths of  $18^\circ$  and  $22.5^\circ$  when measured in degrees). Connection strengths,  $J_{ab}$ , were the same as in Figs 1, 2 and 3 except multiplied by a factor of 1.2. Each neuron in L4 was modeled as a Poisson process with rate given by

$$r_X(\mathbf{x}, \theta) = \bar{r}_X r_{X,x}(\mathbf{x}) r_{X,\theta}(\theta)$$

where  $\mathbf{x}$  is the location of the neuron,  $\theta$  is its preferred orientation,

$$r_{X,x}(\mathbf{x}) = c + (1 - c)G(\mathbf{x} - \mathbf{x}_0; \sigma_X)$$

and

$$r_{X,\theta}(\theta) = c_\theta + (1 - c_\theta)g([\theta - \theta_0]/180^\circ; \sigma_{X,\theta}).$$

This models a stimulus with orientation  $\theta_0 = 0.5$  (representing  $90^\circ$ ) and centered at spatial coordinates  $\mathbf{x}_0 = (0.5, 0.5)$ . The parameters  $\sigma_X$  and  $\sigma_{X,\theta}$  quantify the width of L4 firing rates in physical and orientation space. For all panels in Fig 5, we used  $\sigma_{X,\theta} = 0.1$  (width  $18^\circ$ ) and  $c_\theta = 0.75$ . We used  $\sigma_X = 0.2$  for Fig 5d–5i and  $\sigma_X = 0.06$  for Fig 5j–5o. In both cases, we chose  $\bar{r}_X$  and  $c$  so that the minimum and maximum of  $r_{X,x}(\mathbf{x})$  were 10 and 20 Hz respectively.

For the spatially extended network, the connectivity kernels,  $\mathcal{W}$  and  $\mathcal{W}_X$ , are defined in Results where  $w_{ab}(\mathbf{x}, \theta) = J_{ab} N_b p_{ab}(\mathbf{x}, \theta) / (J_{EXpEX} N_X)$ . The Fourier series in physical and orientation tuning space is defined by

$$\tilde{\mathbf{u}}(\mathbf{n}, k) = \iint \mathbf{u}(\mathbf{x}, \theta) e^{-2\pi i(\mathbf{x}\cdot\mathbf{n} + k\theta)} d\mathbf{x}d\theta$$

where the triple integral is over the two dimensions of physical space and one dimensional

orientation space. The Fourier series of the convolution kernels defined above turns convolution into multiplication in the Fourier domain, from which Eq (10) gives  $\tilde{\mathbf{I}} = (1/\epsilon)[\tilde{\mathbf{W}}\tilde{\mathbf{r}} + \tilde{\mathbf{X}}]$  where  $\tilde{\mathbf{X}}$ ,  $\tilde{\mathbf{W}}$ , and  $\tilde{\mathbf{W}}_X$  are defined in Results with  $\tilde{w}_{ab}(\mathbf{n}, k) = \bar{w}_{ab} \exp[-2\pi^2(|\mathbf{n}|^2\alpha_b^2 + k^2\alpha_{b,\theta}^2)]$ ,  $\bar{w}_{ab} = \tilde{w}_{ab}(\mathbf{0}, 0) = J_{ab}p_{ab}N_b/(J_{EX}p_{EX}N_X)$ , and  $\|\mathbf{n}\|^2 = n_1^2 + n_2^2$ . Using the linear approximation,  $\mathbf{r} = \mathbf{g}\mathbf{I}$  then gives Eq (17). Firing rates for dashed curves in Fig 5 and all firing rates in Figs 6 and 7 were obtained by first computing Eq (17), then inverting the Fourier transform numerically using an inverse fast Fourier transform. Solid curves in Fig 5 were computed similarly, except using Eq (13) in place of Eq (17).

All simulations and numerical computations were performed on a MacBook Pro running OS X 10.9.5 with a 2.3 GHz Intel Core i7 processor. All simulations were written in a combination of C and Matlab (Matlab R 2015b, MathWorks). The differential equations defining the neuron model were solved using a forward Euler method with time step 0.1 ms.

## Supporting information

**S1 Text. Derivation of firing rates in balanced networks with Gaussian-shaped connectivity kernels.** This supplementary text contains details of the analysis of balanced networks with connection probabilities that decay like a Gaussian with distance in physical and orientation space.

(PDF)

**S2 Text. Balance and imbalanced amplification in a model with conductance-based synaptic and optogenetic inputs.** This supplementary text contains simulations and analysis of balance and imbalanced amplification in a network model with conductance-based synapses.

(PDF)

## Acknowledgments

We thank Ashok Litwin-Kumar for helpful comments on a draft of the manuscript.

## Author Contributions

**Conceptualization:** Christopher Ebsch, Robert Rosenbaum.

**Data curation:** Christopher Ebsch, Robert Rosenbaum.

**Formal analysis:** Christopher Ebsch, Robert Rosenbaum.

**Funding acquisition:** Robert Rosenbaum.

**Investigation:** Christopher Ebsch, Robert Rosenbaum.

**Methodology:** Christopher Ebsch, Robert Rosenbaum.

**Software:** Christopher Ebsch, Robert Rosenbaum.

**Supervision:** Robert Rosenbaum.

**Visualization:** Christopher Ebsch, Robert Rosenbaum.

**Writing – original draft:** Christopher Ebsch, Robert Rosenbaum.

**Writing – review & editing:** Christopher Ebsch, Robert Rosenbaum.

## References

1. Tsodyks M, Skaggs W, Sejnowski T, McNaughton B. Paradoxical Effects of External Modulation of Inhibitory Interneurons. *J Neurosci*. 1997; 17(11):4382–4388. PMID: [9151754](#)
2. Ozeki H, Finn IM, Schaffer ES, Miller KD, Ferster D. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*. 2009; 62(4):578–92. <https://doi.org/10.1016/j.neuron.2009.03.028> PMID: [19477158](#)
3. Adesnik H, Scanziani M. Lateral competition for cortical space by layer-specific horizontal circuits. *Nature*. 2010; 464(7292):1155–60. <https://doi.org/10.1038/nature08935> PMID: [20414303](#)
4. Phillips E, Hasenstaub A. Asymmetric effects of activating and inactivating cortical interneurons. *eLife*. 2016; 5:e18383. <https://doi.org/10.7554/eLife.18383> PMID: [27719761](#)
5. Chambers B, MacLean JN. Higher-Order Synaptic Interactions Coordinate Dynamics in Recurrent Networks. *PLoS Comput Biol*. 2016; 12(8):e1005078. <https://doi.org/10.1371/journal.pcbi.1005078> PMID: [27542093](#)
6. van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*. 1996; 274(5293):1724–1726. <https://doi.org/10.1126/science.274.5293.1724> PMID: [8939866](#)
7. van Vreeswijk C, Sompolinsky H. Chaotic balanced state in a model of cortical circuits. *Neural Comput*. 1998; 10(6):1321–1371. <https://doi.org/10.1162/089976698300017214> PMID: [9698348](#)
8. Rosenbaum R, Doiron B. Balanced networks of spiking neurons with spatially dependent recurrent connections. *Phys Rev X*. 2014; 4(2):021039.
9. Pyle R, Rosenbaum R. Highly connected neurons spike less frequently in balanced networks. *Phys Rev E*. 2016; 93(4):040302. <https://doi.org/10.1103/PhysRevE.93.040302> PMID: [27176240](#)
10. Landau ID, Egger R, Dercksen VJ, Oberlaender M, Sompolinsky H. The Impact of Structural Heterogeneity on Excitation-Inhibition Balance in Cortical Networks. *Neuron*. 2016;.
11. Renart A, de La Rocha J, Bartho P, Hollender L, Parga N, Reyes A, et al. The Asynchronous State in Cortical Circuits. *Science*. 2010; 327(5965):587–590. <https://doi.org/10.1126/science.1179850> PMID: [20110507](#)
12. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B. The spatial structure of correlated neuronal variability. *Nature Neurosci*. 2016;.
13. Haider B, Duque A, Hasenstaub AR, McCormick DA. Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J Neurosci*. 2006; 26(17):4535–4545. <https://doi.org/10.1523/JNEUROSCI.5297-05.2006> PMID: [16641233](#)
14. Okun M, Lampl I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neurosci*. 2008; 11(5):535–537. <https://doi.org/10.1038/nn.2105> PMID: [18376400](#)
15. Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. *Nature Neurosci*. 2011; 14(7):811–819. <https://doi.org/10.1038/nn.2842> PMID: [21709677](#)
16. Cohen-Kashi Malina K, Jubran M, Katz Y, Lampl I. Imbalance between excitation and inhibition in the somatosensory cortex produces postadaptation facilitation. *J Neurosci*. 2013; 33(19):8463–8471. <https://doi.org/10.1523/JNEUROSCI.4845-12.2013> PMID: [23658183](#)
17. Marlin BJ, Mitre M, D'amour JA, Chao MV, Froemke RC. Oxytocin enables maternal behaviour by balancing cortical inhibition. *Nature*. 2015; 520(7548):499–504. <https://doi.org/10.1038/nature14402> PMID: [25874674](#)
18. Dehghani N, Peyrache A, Telenczuk B, Le Van Quyen M, Halgren E, Cash SS, et al. Dynamic Balance of Excitation and Inhibition in Human and Monkey Neocortex. *Sci Rep*. 2016; 6. <https://doi.org/10.1038/srep23176>
19. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K. The mechanics of state-dependent neural correlations. *Nature Neurosci*. 2016; 19(3):383–393. <https://doi.org/10.1038/nn.4242> PMID: [26906505](#)
20. Gerstein GL, Mandelbrot B. Random Walk Models for the Spike Activity of a Single Neuron. *Biophys J*. 1964; 4(c):41–68.
21. Softky WR, Koch C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci*. 1993; 13(1):334–50. PMID: [8423479](#)
22. Amit D, Brunel N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex*. 1997; 7(3):237–252. <https://doi.org/10.1093/cercor/7.3.237> PMID: [9143444](#)
23. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci*. 1998; 18(10):3870–96. PMID: [9570816](#)



24. Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* 2005; 8(9):1263–1268. <https://doi.org/10.1038/nn1525> PMID: 16116447
25. Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience. *Nature Neurosci.* 2015; 18(9):1213–1225. <https://doi.org/10.1038/nn.4091> PMID: 26308982
26. Brette R, Gerstner W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J Neurophysiol.* 2005; 94(5):3637–3642. <https://doi.org/10.1152/jn.00686.2005> PMID: 16014787
27. Jolivet R, Rauch A, Lüscher HR, Gerstner W. Integrate-and-Fire models with adaptation are good enough: predicting spike times under random current injection. *Adv Neural Inf Process Syst.* 2006; 18:595–602.
28. Jolivet R, Schürmann F, Berger TK, Naud R, Gerstner W, Roth A. The quantitative single-neuron modeling competition. *Biol Cybern.* 2008; 99(4-5):417–426. <https://doi.org/10.1007/s00422-008-0261-x> PMID: 19011928
29. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci.* 1998; 18(10):3870–3896. PMID: 9570816
30. Ecker A, Berens P, Keliris G, Bethge M, Logothetis N, Tolias A. Decorrelated Neuronal Firing in Cortical Microcircuits. *Science.* 2010; 327(5965):584–587. <https://doi.org/10.1126/science.1179867> PMID: 20110506
31. Brunel N, Hakim V. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.* 1999; 11(7):1621–1671. <https://doi.org/10.1162/089976699300016179> PMID: 10490941
32. Renart A, Brunel N, Wang XJ. Mean-field theory of irregularly spiking neuronal populations and working memory in recurrent cortical networks. In: *Computational Neuroscience: A Comprehensive Approach.* CRC Press, New York; 2004. p. 431–490.
33. Ledoux E, Brunel N. Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. *Front Comput Neurosci.* 2011; 5:25. <https://doi.org/10.3389/fncom.2011.00025> PMID: 21647353
34. Tan A, Chen Y, Scholl B, Seidemann E, Priebe NJ. Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature.* 2014; 509(7499):226. <https://doi.org/10.1038/nature13159> PMID: 24695217
35. Dayan P, Abbott LF. *Theoretical Neuroscience.* Cambridge, MA: MIT Press; 2001.
36. Hertäg L, Durstewitz D, Brunel N. Analytical approximations of the firing rate of an adaptive exponential integrate-and-fire neuron in the presence of synaptic noise. *Frontiers in Comput Neurosci.* 2014; 8.
37. Petreanu L, Huber D, Sobczyk A, Svoboda K. Channelrhodopsin-2–assisted circuit mapping of long-range callosal projections. *Nature Neurosci.* 2007; 10(5):663–668. <https://doi.org/10.1038/nn1891> PMID: 17435752
38. Pouille F, Marin-Burgin A, Adesnik H, Atallah BV, Scanziani M. Input normalization by global feedforward inhibition expands cortical dynamic range. *Nature Neurosci.* 2009; 12(12):1577–1585. <https://doi.org/10.1038/nn.2441> PMID: 19881502
39. Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J. The dynamics of sensory integration in a hierarchical network explains choice probabilities in MT. *Nature Commun.* 2015; 6:1–13.
40. Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, et al. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neurosci.* 2011; 14(8):1045–52. <https://doi.org/10.1038/nn.2876> PMID: 21765421
41. Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD. Functional specificity of local synaptic connections in neocortical networks. *Nature.* 2011; 473(7345):87–91. <https://doi.org/10.1038/nature09880> PMID: 21478872
42. Levy RB, Reyes AD. Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. *J Neurosci.* 2012; 32(16):5609–5619. <https://doi.org/10.1523/JNEUROSCI.5158-11.2012> PMID: 22514322
43. Lim S, Goldman MS. Balanced cortical microcircuitry for spatial working memory based on corrective feedback control. *J Neurosci.* 2014; 34(20):6790–6806. <https://doi.org/10.1523/JNEUROSCI.4602-13.2014> PMID: 24828633
44. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015; 521(7553):436–444. <https://doi.org/10.1038/nature14539> PMID: 26017442
45. Tricomi FG. *Integral equations.* Interscience, New York; 1957.

46. Ben-Yishai R, Bar-Or RL, Sompolinsky H. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci USA*. 1995; 92(9):3844–3848. <https://doi.org/10.1073/pnas.92.9.3844> PMID: 7731993
47. Rubin DB, Van Hooser SD, Miller KD. The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*. 2015; 85(2):1–51. <https://doi.org/10.1016/j.neuron.2014.12.026>
48. Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M. A neural circuit for spatial summation in visual cortex. *Nature*. 2012; 490(7419):226–31. <https://doi.org/10.1038/nature11526> PMID: 23060193
49. Jiang X, Shen S, Cadwell CR, Berens P, Sinz F, Ecker AS, et al. Principles of connectivity among morphologically defined cell types in adult neocortex. *Science*. 2015; 350(6264):aac9462. <https://doi.org/10.1126/science.aac9462> PMID: 26612957
50. Ermentrout B. Linearization of F-I curves by adaptation. *Neural Comput*. 1998; 10(7):1721–1729. <https://doi.org/10.1162/089976698300017106> PMID: 9744894
51. Koch Ocker G, Doiron B. Kv7 channels regulate pairwise spiking covariability in health and disease. *J Neurophysiol*. 2014; 112(2):340–352. <https://doi.org/10.1152/jn.00084.2014>
52. Miller KD. Canonical computations of cerebral cortex. *Curr Opin Neurobiol*. 2016; 37:75–84. <https://doi.org/10.1016/j.conb.2016.01.008> PMID: 26868041
53. Litwin-Kumar A, Rosenbaum R, Doiron B. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. *J Neurophysiol*. 2016; 115(3):1399–1409. <https://doi.org/10.1152/jn.00732.2015> PMID: 26740531
54. Atallah B, Bruns W, Carandini M, Scanziani M. Parvalbumin-Expressing Interneurons Linearly Transform Cortical Responses to Visual Stimuli. *Neuron*. 2012; 73(1):159–170. <https://doi.org/10.1016/j.neuron.2011.12.013> PMID: 22243754
55. Murphy BK, Miller KD. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*. 2009; 61(4):635–48. <https://doi.org/10.1016/j.neuron.2009.02.005> PMID: 19249282
56. Bosking WH, Zhang Y, Schofield B, Fitzpatrick D. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J Neurosci*. 1997; 17(6):2112–27. PMID: 9045738
57. Destexhe A, Paré D. Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *J Neurophysiol*. 1999; 81(4):1531–47. <https://doi.org/10.1152/jn.1999.81.4.1531> PMID: 10200189
58. Destexhe A, Rudolph M, Paré D. The high-conductance state of neocortical neurons in vivo. *Nat Rev Neurosci*. 2003; 4(4):739–51. <https://doi.org/10.1038/nrn1198> PMID: 12951566
59. Kuhn A, Aertsen A, Rotter S. Neuronal integration of synaptic input in the fluctuation-driven regime. *J Neurosci*. 2004; 24(10):2345–56. <https://doi.org/10.1523/JNEUROSCI.3349-03.2004> PMID: 15014109
60. Kameda H, Hioki H, Tanaka YH, Tanaka T, Sohn J, Sonomura T, et al. Parvalbumin-producing cortical interneurons receive inhibitory inputs on proximal portions and cortical excitatory inputs on distal dendrites. *Europ J Neurosci*. 2012; 35(6):838–854. <https://doi.org/10.1111/j.1460-9568.2012.08027.x>
61. Hioki H, Okamoto S, Konno M, Kameda H, Sohn J, Kuramoto E, et al. Cell type-specific inhibitory inputs to dendritic and somatic compartments of parvalbumin-expressing neocortical interneuron. *J Neurosci*. 2013; 33(2):544–555. <https://doi.org/10.1523/JNEUROSCI.2255-12.2013> PMID: 23303934
62. Litwin-Kumar A, Doiron B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neurosci*. 2012; 15(11):1498–1505. <https://doi.org/10.1038/nn.3220> PMID: 23001062
63. Stringer C, Pachitariu M, Steinmetz NA, Okun M, Bartho P, Harris KD, et al. Inhibitory control of correlated intrinsic variability in cortical networks. *eLife*. 2016; 5:e19695. <https://doi.org/10.7554/eLife.19695> PMID: 27926356