RESEARCH

Check for updates

Spatially extended balanced networks without translationally invariant connectivity

Christopher Ebsch[1] and Robert Rosenbaum[1,2*] (iD)

*Correspondence:
Robert.Rosenbaum@nd.edu
[1]Department of Applied and
Computational Mathematics and
Statistics, University of Notre Dame,
Notre Dame, USA
[2]Interdisciplinary Center for
Network Science and Applications,
University of Notre Dame, Notre
Dame, USA

Abstract

Networks of neurons in the cerebral cortex exhibit a balance between excitation (positive input current) and inhibition (negative input current). Balanced network theory provides a parsimonious mathematical model of this excitatory-inhibitory balance using randomly connected networks of model neurons in which balance is realized as a stable fixed point of network dynamics in the limit of large network size. Balanced network theory reproduces many salient features of cortical network dynamics such as asynchronous-irregular spiking activity. Early studies of balanced networks did not account for the spatial topology of cortical networks. Later works introduced spatial connectivity structure, but were restricted to networks with translationally invariant connectivity structure in which connection probability depends on distance alone and boundaries are assumed to be periodic. Spatial connectivity structure in cortical network does not always satisfy these assumptions. We use the mathematical theory of integral equations to extend the mean-field theory of balanced networks to account for more general dependence of connection probability on the spatial location of pre- and postsynaptic neurons. We compare our mathematical derivations to simulations of large networks of recurrently connected spiking neuron models.

Keywords: Balanced networks; Spiking neural network models; Excitatory-inhibitory balance; Mean-field theory

1 Introduction

Balanced networks [1, 2] offer a parsimonious computational and mathematical model of the asynchronous-irregular spiking activity and excitatory-inhibitory balance that are ubiquitous in cortical neuronal networks [3–10]. Balanced networks can produce asynchronous and irregular activity through chaotic or chaos-like spike timing dynamics [2, 11]. Mean-field analysis of balanced networks reveals a stable fixed point that naturally produces excitatory-inhibitory balance and weak pairwise spike train correlations without fine tuning of model parameters. In early studies, this mean-field analysis was performed in networks in which connection probabilities are homogeneous across the excitatory and inhibitory populations [1]. Later work extended the mean-field analysis to networks with multiple sub-populations and to spatially extended networks with distance-dependent connection probabilities [12–16], including models that combine physical and

Springer

tuning space [17]. Previous work on spatially extended balanced networks assumed that connection probabilities depended only on the distance between neurons measured with periodic boundaries, rendering connection probabilities translationally invariant. This assumption allows the Fourier modes of network activity to decouple, so the mean-field equations can be easily solved in the Fourier domain. However, connectivity in cortical neuronal networks is not so simple. While the use of periodic boundaries is justified for modeling naturally periodic spaces like orientation tuning space, it is not necessarily realistic for models of physical space. Moreover, connection probabilities in cortical neuronal networks can depend on neuron location in more complicated ways than a pure distance dependence [18–20].

We use the mathematical theory of integral equations [21] to extend the mean-field theory of firing rates in balanced networks, permitting a more general dependence of connection probability on the spatial location of pre- and postsynaptic neurons. We derive conditions on the spatial structure connectivity and external input under which networks can maintain balance in the limit of large network size, derive the spatial profile of firing rates in this limit when balance is maintained, and derive a linear approximation to firing rates when balance is broken. We demonstrate our findings with simulations of large spiking networks under a simple spatial connectivity structure that violates the translational invariance of connection probabilities assumed by previous work.

2 Model and background

We consider recurrent networks of N model neurons, $N_e = 0.8N$ of which are excitatory and $N_i = 0.2N$ inhibitory. The membrane potential of neuron j in population $a = e, i$ obeys the exponential integrate-and-fire dynamics

$$C_m \frac{dV_j^a}{dt} = -g_L\left(V_j^a - E_L\right) + g_L D_T e^{(V_j^a - V_T)/D_T} + I_j^a(t)$$

with the added condition that each time $V_j^a(t)$ exceeds V_{th}, it is reset to V_{re}, held for a refractory period of τ_{ref}, and a spike is recorded at the threshold crossing time. The synaptic input current to neuron j in population a is given by

$$I_j^a(t) = \sum_{b=e,i} \sum_{k=1}^{N_b} J_{jk}^{ab} \alpha_b\left(t - t_n^{a,k}\right) + \sqrt{N} F_j^a, \tag{1}$$

where $t_n^{a,j}$ is the nth spike time of neuron j in population $a = e, i$ and $\alpha_b(t) = (1/\tau_b)e^{-t/\tau_b}H(t)$ is a postsynaptic current waveform, and $H(t)$ is the Heaviside step function.

We consider a network on the compact domain $[0, 1]$ with neuron $j = 1, \dots, N_a$ in population $a = e, i$ located at $x = j/N_a$. Connection strengths are defined by

$$J_{jk}^{ab} = \begin{cases} j_{ab}/\sqrt{N} & \text{with probability } p_{ab}(x, y), \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where $x = j/N_a \in \Omega$ is the location of (postsynaptic) neuron j in population a and $y = k/N_b \in \Omega$ is the location of (presynaptic) neuron k in population b. We assume that $j_{ae} > 0$ and $j_{ai} < 0$ with all j_{ab} constant with respect to N. The model and mean-field theory

are easily extended to the case where connection strength and neuron density are spatially inhomogeneous (see below). The $1/\sqrt{N}$ scaling of synaptic weights is a defining feature of balanced networks which naturally captures excitatory-inhibitory balance and asynchronous-irregular spiking activity observed in cortical recordings [1, 2, 22]. Recent work in cultured cortical populations shows that similar scaling laws emerge naturally and produce network dynamics consistent with the balanced state [23]. Feedforward external input to the network is modeled by

$$F_j^a = F_a(x),$$

where $x \in \Omega$ is the location of neuron j in population a. This models the input from $\mathcal{O}(N)$ neurons with synaptic weights that scale like $\mathcal{O}(1/\sqrt{N})$. Indeed, external input can be replaced by a population of generated spike trains without affecting the mean-field theory [17].

We consider a simple example in which $\Omega = [0, 1]$ and

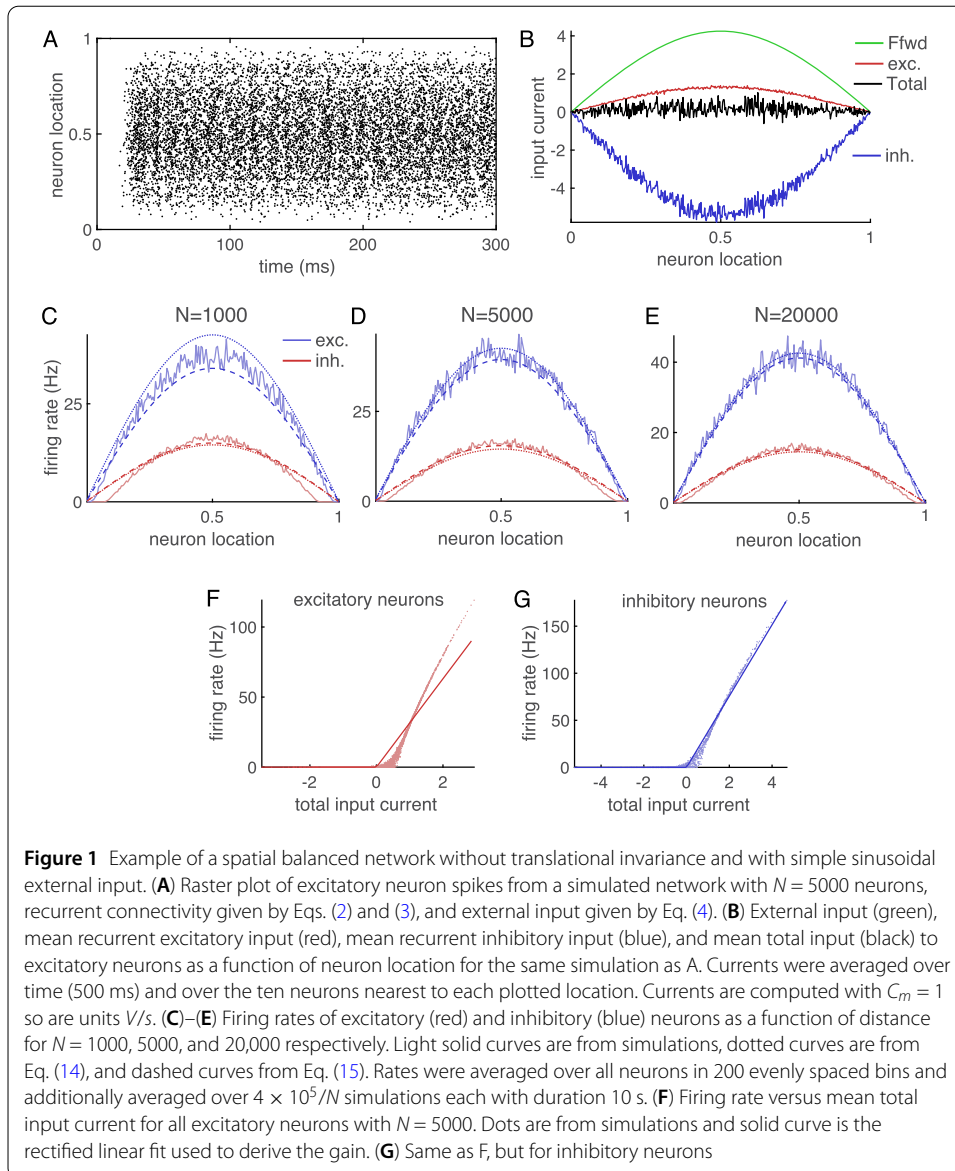$$p_{ab}(x, y) = 12\overline{p_{ab}}\big(\min(x, y) - xy\big), \tag{3}$$

where $\overline{p_{ab}} = \iint p_{ab}(x, y)\, dx\, dy$ is the network-averaged connection probability from b to a. Note that we use the convention that p_{ab}, j_{ab}, etc. refer to connection probabilities and strengths from presynaptic population b to postsynaptic population a. Note that, unlike spatial balanced networks considered in previous work [12, 17], $p_{ab}(x, y)$ is not translationally invariant and cannot be written in terms of $x - y$. Specifically, neurons near the center of the domain send and receive more connections than neurons near the edges.

We first simulated this network with $N = 5000$ neurons and

$$F_a(x) = \overline{F}_a \sin(\pi x), \tag{4}$$

where $\overline{F}_a > 0$. Our simulation produced asynchronous, irregular spiking activity (Fig. 1(A)) and excitatory-inhibitory balance (Fig. 1(B)) that are characteristic of the balanced network state [1, 2] and of cortical neuronal networks [3–10]. Note that the simulations are entirely deterministic once the random network connectivity and initial conditions are specified, so irregular spiking is driven by chaotic or chaos-like dynamics [2, 11], not noise. Firing rates were peaked near the center of the domain and decayed toward zero near the edge for various values of N (Fig. 1(C)–(E)). Firing rates that are peaked at the center of the spatial domain are not unexpected, given the structure of our connectivity kernel and input, and are common across many models of spatially extended networks [12, 24, 25]. We next derive a mean-field theory for computing firing rates in spatially extended balanced networks like this one.

Parameters for all figures are $C_m/g_L = 15$ ms, $E_L = -72$ mV, $V_T = -60$ mV, $V_{th} = -15$ mV, $V_{re} = -72$, $V_{lb} = -100$, $\tau_{ref} = 1$ ms, $\Delta_T = 1.5$ mV, $\tau_e = 8$ ms, $\tau_i = 4$ ms, $j_{ee}/C_m = 25$ mV, $j_{ei}/C_m = -150$ mV, $j_{ie}/C_m = 112.5$ mV, $j_{ii}/C_m = -250$ mV, $\overline{F}_e/C_m = 0.06$ V/s, $\overline{F}_i = 0.05$ V/s, and $\overline{p}_{ee} = \overline{p}_{ei} = \overline{p}_{ie} = \overline{p}_{ii} = 0.05$. Matlab and C code for running all simulations and generating all figures is publicly available at https://github.com/RobertRosenbaum/SpatialBalNetNonTransInvar/.

Figure 1 Example of a spatial balanced network without translational invariance and with simple sinusoidal external input. (**A**) Raster plot of excitatory neuron spikes from a simulated network with $N = 5000$ neurons, recurrent connectivity given by Eqs. (2) and (3), and external input given by Eq. (4). (**B**) External input (green), mean recurrent excitatory input (red), mean recurrent inhibitory input (blue), and mean total input (black) to excitatory neurons as a function of neuron location for the same simulation as A. Currents were averaged over time (500 ms) and over the ten neurons nearest to each plotted location. Currents are computed with $C_m = 1$ so are units *V/s*. (**C**)–(**E**) Firing rates of excitatory (red) and inhibitory (blue) neurons as a function of distance for $N = 1000$, 5000, and 20,000 respectively. Light solid curves are from simulations, dotted curves are from Eq. (14), and dashed curves from Eq. (15). Rates were averaged over all neurons in 200 evenly spaced bins and additionally averaged over $4 \times 10^5/N$ simulations each with duration 10 s. (**F**) Firing rate versus mean total input current for all excitatory neurons with $N = 5000$. Dots are from simulations and solid curve is the rectified linear fit used to derive the gain. (**G**) Same as F, but for inhibitory neurons

3 Mean field theory of balance in spatially extended networks without translational invariance

The mean-field theory is developed by considering the large N limit and defining the mean-field firing rates

$$\boldsymbol{r}(x) = \begin{bmatrix} r_e(x) \\ r_i(x) \end{bmatrix},$$

where $r_a(x)$ is the expected value of the firing rates of a neuron in population a at location $x \in \Omega$. The mean total input $\boldsymbol{I}(x) = [I_e(x)I_i(x)]^T$ and the feedforward external input $\boldsymbol{F}(x) = [F_e(x)F_i(x)]^T$ are defined analogously. We assume that $F_a \in L^2(\Omega)$.

The following equation gives a heuristic approximation to neurons' input as a function of firing rates in the network [12, 17]:

$$\boldsymbol{I} = \sqrt{N}[\mathcal{W}\boldsymbol{r} + \boldsymbol{F}]. \tag{5}$$

Here, $\mathcal{W}$ is an integral operator defined by

$$\mathcal{W}\boldsymbol{r} = \begin{bmatrix} \mathcal{W}_{ee}r_e + \mathcal{W}_{ei}r_i \\ \mathcal{W}_{ie}r_e + \mathcal{W}_{ii}r_i \end{bmatrix}, \tag{6}$$

where

$$[\mathcal{W}_{ab}r_b](x) = \int_\Omega w_{ab}(x,y)r_b(y)\,dy, \tag{7}$$

$w_{ab}(x,y) = p_{ab}(x,y)j_{ab}q_b \sim \mathcal{O}(1)$ is the mean-field connectivity kernel with $q_e = N_e/N = 0.8$, and $q_i = N_i/N = 0.2$. We assume that $w_{ab} \in L^2(\Omega^2)$ so that $\mathcal{W}_{ab}$ and $\mathcal{W}$ are Hilbert–Schmidt integral operators and are, therefore, compact [21]. We further assume that $w_{ab}(x,y) = w_{ab}(y,x)$ so that $\mathcal{W}_{ab}$ is a Hermitian integral operator (see Discussion for comments on relaxing some of these assumptions).

Note that $\mathcal{W}$ and $\boldsymbol{F}$ do not depend on N, but $\boldsymbol{r}$ and $\boldsymbol{I}$ do depend on N. In the balanced network theory below, we derive expressions for $\boldsymbol{r}$ that must be satisfied for $\|\boldsymbol{r}\|$ and $\|\boldsymbol{I}\|$ to be finite in the $N \to \infty$ limit (where $\|\cdot\|$ is the L^2 norm).

Equation (5) quantifies how firing rates are mapped to mean synaptic input in the network. A closed form approximation in Eq. (5) is possible in this case because the mapping is linear. However, the mapping from synaptic input to firing rates is necessarily nonlinear, depends on the details of the neuron model, can depend on higher moments of the input currents, and is generally not amenable to a closed-form mathematical approximation for the spiking network model considered here. Some studies use a diffusion approximation and Fokker–Planck formalism to account for the dependence of neurons' firing rates on the first two moments of their input currents [26–28]. This approach can yield accurate approximations in practice, but makes the assumption that synaptic input is accurately approximated by Gaussian white noise, which may be inaccurate in some settings.

The mean-field theory of balanced networks offers an alternative approach to analyzing firing rates in which the mapping from synaptic input statistics to rates does not need to be known. This theory is developed and applied by analyzing Eq. (5) and the integral equations implied by it, which serves as a heuristic approximation to our spiking network model. We then compare the results of our analysis to simulations of the spiking network model. For balanced network theory, we do not need to specify the exact mapping from $\boldsymbol{I}$ to $\boldsymbol{r}$. Instead, the only necessary condition for the analysis of balanced networks is that the mapping does not converge to zero or diverge with N, more specifically that

$$\boldsymbol{r}, \boldsymbol{I} \sim \mathcal{O}(1) \tag{8}$$

as $N \to \infty$, where all orders of magnitude expressions (expressions of the form $U \sim \mathcal{O}(F(N))$) should be interpreted in the $N \to \infty$ limit under the L^2 norm, so the expression above means $\lim_{N\to\infty} \|\boldsymbol{r}\|, \|\boldsymbol{I}\| < \infty$, where $\|\cdot\|$ is the L^2 norm on Ω. From Eq. (5), one

can see that satisfying Eq. (8) implies that the network produces large $\mathcal{O}(\sqrt{N})$, excitatory and inhibitory synaptic currents that cancel or "balance" each other to produce $\mathcal{O}(1)$ total synaptic input. Hence, networks that satisfy Eq. (8) condition are said to be "balanced networks" or to operate in a "balanced state."

Remarkably, this condition alone is enough to derive a linear, closed form expression for firing rates in the $N \to \infty$ limit even if the mapping from $\boldsymbol{I}$ to $\boldsymbol{r}$ is unknown. To see this, note that Eqs. (5) and (8) can only be realized when firing rates satisfy

$$\mathcal{W}\boldsymbol{r} + \boldsymbol{F} = 0 \tag{9}$$

in the $N \to \infty$ limit. In other words, if $\boldsymbol{r}_\infty = \lim_{N \to \infty} \boldsymbol{r}$ exists and is finite, then it must satisfy $\mathcal{W}\boldsymbol{r}_\infty + \boldsymbol{F} = 0$. Note that $\mathcal{W}$ and $\boldsymbol{F}$ do not depend on N. Hence, firing rates in the $N \to \infty$ limit in balanced networks are determined by a linear integral equation of the first kind [21] despite the nonlinearity of the mapping from $\boldsymbol{I}$ to $\boldsymbol{r}$.

To better understand how this works, consider a commonly used integro-differential equation model for firing rates [24, 25] $\tau \dot{\boldsymbol{r}} = -\boldsymbol{r} + g(\boldsymbol{I})$, where g is a monotonically increasing function. Any fixed point of this system satisfies $\mathcal{W}\boldsymbol{r} + \boldsymbol{F} = g^{-1}(\boldsymbol{r})/\sqrt{N}$. As long as $g^{-1}(\boldsymbol{r})/\sqrt{N} \to 0$ as $N \to \infty$ at the fixed point, any fixed point approaches the solution to Eq. (9) as $N \to \infty$.

Since Eq. (9) is a Fredholm integral equation of the first kind, it does not generically admit a solution. More specifically, for any integral operator $\mathcal{W}$, there necessarily exist external input profiles $\boldsymbol{F}(x)$, for which there is no solution $\boldsymbol{r}(x)$ to Eq. (9). When this occurs, the network cannot satisfy Eq. (8). Therefore, any spatially extended network can be imbalanced by some external input profiles. Moreover, solutions to Eq. (9) must be nonnegative and stable for the balanced state to be realized (see Discussion).

When Eq. (9) does not admit a solution, a linear approximation provides an approximation to firing rates at finite N [12, 17]. This approximation is obtained by making a rectified linear approximation to the mapping from mean input to mean firing rates $r_a(x) = g_a[I_a(x)]_+$ for $a = e, i$, where $g_a > 0$ is the neurons' gain and $[\cdot]_+$ denotes the positive part. For spiking network models, g_a can be approximated heuristically [12] or by fitting simulation results to a rectified linear function [17]. In the examples considered below, we use the latter approach. When firing rates are positive, this gives the following integral equation for firing rates:

$$\mathcal{W}\boldsymbol{r} + \boldsymbol{F} = \epsilon D \boldsymbol{r}, \tag{10}$$

where $\epsilon = 1/\sqrt{N}$ is a small, positive number and

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

This equation is a finite N correction to Eq. (9). Since it is an integral equation of the second kind, it generically admits a unique solution for any $\boldsymbol{F}$ (unless ϵ is an eigenvalue of $\mathcal{W}$). Therefore, even when Eq. (9) does not admit a solution, Eq. (10) will generally admit a solution. However, the solution to Eq. (10) can diverge as $N \to \infty$ (*i.e.*, $\epsilon \to 0$), indicating networks for which Eq. (9) does not admit a solution.

A common approach to solving Fredholm integral equations like Eqs. (9) and (10) is to expand the equations using an orthonormal basis of eigenfunctions for the integral operator [21]. However, the integral operator is not guaranteed to have orthogonal eigenfunctions if it is not Hermitian. Because we assume $w_{ab}(x, y) = w_{ab}(y, x)$, the integral operators $\mathcal{W}_{ab}$ are Hermitian. However, the integral operator $\mathcal{W}$ that comprises our integral equations is not Hermitian because $w_{ei} \neq w_{ie}$. Hence, even though $\mathcal{W}_{ab}$ have orthogonal eigenfunctions, $\mathcal{W}$ does not. We extended the standard theory of integral equations to account for this case in which a non-Hermitian integral operator is composed of multiple Hermitian operators. Our extension is summarized by the following theorem. All convergences in the theorem and proof should be interpreted in an L^2 sense.

Theorem 1 *Suppose that $\mathcal{W}$ is defined as in Eqs. (6) and (7) where $w_{ab}(x, y) = w_{ab}(y, x)$, $w_{ab} \in L^2(\Omega^2)$, and all four operators $\mathcal{W}_{ab}$ share the same orthonormal basis of eigenfunctions $\{\phi_m \in L^2(\Omega)\}_m$ with associated eigenvalues $\{\mu_m^{ab}\}_m$. When $\epsilon \neq 0$ and $\epsilon D - \widetilde{W}_m$ is nonsingular for all m,*

$$\boldsymbol{r}(x) = \sum_m [\epsilon D - \widetilde{W}_m]^{-1} \widetilde{\boldsymbol{F}}_m \phi_m(x) \tag{11}$$

converges to a solution to Eq. (10) where

$$\widetilde{W}_m = \begin{bmatrix} \mu_m^{ee} & \mu_m^{ei} \\ \mu_m^{ie} & \mu_m^{ii} \end{bmatrix}, \qquad \widetilde{\boldsymbol{F}}_m = \begin{bmatrix} \langle F_e, \phi_m \rangle \\ \langle F_i, \phi_m \rangle \end{bmatrix},$$

and $\langle \cdot, \cdot \rangle$ is the L^2 inner product on Ω. When the series in Eq. (11) converges at $\epsilon = 0$, it converges to a solution to Eq. (9).

Proof All convergences should be interpreted in the L^2 sense. Note that, since $w_{ab}(x, y) = w_{ab}(y, x)$ and $w_{ab} \in L^2$, $\mathcal{W}_{ab}$ is a self-adjoint Hilbert–Schmidt integral operator and is therefore a compact operator; therefore $\{\phi_m\}_m$ is a complete basis for $L^2(\Omega)$ and $\mu_m^{ab} \in \mathbb{R}$ by the spectral theorem.

We first show that the series converges when $\epsilon \neq 0$. Note that $\mu_m^{ab} \to 0$ as $m \to \infty$ by the completeness of $\{\phi_m\}$ so that $\epsilon D - \widetilde{W}_m \to \epsilon D$ as $m \to \infty$ when $\epsilon \neq 0$. Therefore, the series in Eq. (11) converges only if the series $\sum_m [\epsilon D]^{-1} \widetilde{\boldsymbol{F}}_m \phi_m$ converges. The completeness of $\{\phi_m\}$ implies that $\sum_m \widetilde{\boldsymbol{F}}_m \phi_m$ converges since $\sum_m \langle F_a, \phi_m \rangle \phi_m$ converges. We may conclude that the series in Eq. (11) converges.

We next show that the series in Eq. (11) converges to a solution to Eq. (10). Let $\boldsymbol{r}(x) = [r_e(x) r_i(x)]^T$ be defined by Eq. (11). Then

$$\langle \boldsymbol{r}, \phi_m \rangle \equiv \begin{bmatrix} \langle r_e, \phi_m \rangle \\ \langle r_i, \phi_m \rangle \end{bmatrix} = [\epsilon D - \widetilde{W}_m]^{-1} \widetilde{\boldsymbol{F}}_m.$$

Multiplying both sides by $[\epsilon D - \widetilde{W}_m]$ gives

$$\epsilon D \langle \boldsymbol{r}, \phi_m \rangle - \widetilde{W}_m \langle \boldsymbol{r}, \phi_m \rangle = \widetilde{\boldsymbol{F}}_m,$$

which gives

$$\langle \epsilon D \boldsymbol{r}, \phi_m \rangle - \langle \mathcal{W} \boldsymbol{r}, \phi_m \rangle = \langle \boldsymbol{F}, \phi_m \rangle$$

for all m. Since ϕ_m is a complete basis, this implies that $\epsilon D\mathbf{r} - \mathcal{W}\mathbf{r} = \mathbf{F}$ and therefore $\mathbf{r}$ satisfies Eq. (10).

Now assume that the series in Eq. (11) converges at $\epsilon = 0$. Repeating the argument above with $\epsilon = 0$ shows that it converges to a solution to Eq. (9) whenever it converges. □

Note that the convergence of Eq. (11) when $\epsilon > 0$ is implied by the assumptions of the theorem, but the convergence when $\epsilon = 0$ needs to be assumed separately. Hence, Eq. (10) admits solutions under relatively general assumptions, whereas the solvability of Eq. (9) is less general and requires that $\|\widetilde{\mathbf{F}}_m\|$ converges to zero faster than $\|\widetilde{W}_m\|$.

While the assumption of symmetric convolution kernels assures that $\mathcal{W}_{ab}$ have orthonormal bases of eigenfunctions with real eigenvalues, we additionally assumed that these eigenfunctions were the same for all four combinations of $a, b \in \{e, i\}$. This would be satisfied, for example, if $w_{ab}(x, y) = \overline{w}_{ab}k(x, y)$ for some $\overline{w}_{ab} \in \mathbb{R}$, but also applies to other settings. For example, if $w_{ab}(x, y) = w_{ab}(x - y)$ are periodic or "wrapped" convolution kernels as in previous work [12, 15, 17], then the Fourier basis functions provide an orthonormal basis of eigenfunctions even when the four convolution kernels $w_{ab}(x-y)$ are not multiples of each other. In this case, Eq. (11) is the Fourier series for $\mathbf{r}(x)$. In Discussion, we comment on how some of these assumptions could be weakened.

When $\mathcal{W}_{ab}$ do not share orthonormal bases of eigenfunctions, the analysis must be performed directly on the spectrum of $\mathcal{W}$ instead of decomposing the analysis into the spectra of each $\mathcal{W}_{ab}$. Specifically, if $\mathcal{W}$, treated as an integral operator on $L^2(\Omega) \times L^2(\Omega)$, has an orthonormal basis of eigenfunctions $\{\phi_m\}_m$ where $\phi_m \in L^2(\Omega) \times L^2(\Omega)$ with real eigenvalues $\{\mu_m\}$, then $\mathbf{r} = \sum_m [\epsilon D - \mu_m]^{-1}\langle \mathbf{F}, \phi_m\rangle\phi_m$. However, note that $\mathcal{W}$ may not have an orthonormal basis of eigenfunctions with real eigenvalues even when kernels are spatially symmetric ($w_{ab}(x, y) = w_{ab}(y, x)$) because a Hermitian $\mathcal{W}$ would also require that $w_{ei}(x, y) = w_{ie}(y, x)$, which is generally not true for excitatory-inhibitory networks. Solving Fredholm equations with non-Hermitian integral operators is generally more difficult as diagonalization methods cannot be applied directly. Hence, our derivation of Eq. (11) extends standard diagonalization methods because it solves a Fredholm equation with a non-Hermitian kernel, albeit one with a special structure in which $\mathcal{W}$ is composed of multiple Hermitian kernels. This structure arises naturally in spatial neural network models with excitatory and inhibitory populations.

A corollary gives a simpler expression for the solution to Eq. (9) when the spatial shape of the connectivity kernels and external input are the same for excitatory and inhibitory neurons.

Corollary 1 *Suppose, in addition to the assumptions of Theorem* 1, *that* $w_{ab}(x, y) = \overline{w}_{ab}k(x, y)$ *and* $F_a(x) = \overline{F}_a F(x)$ *for some* $\overline{w}_{ab}, \overline{F}_a \in \mathbb{R}$. *Then a solution to Eq.* (9) *exists and is equal to*

$$\mathbf{r}(x) = -\overline{W}^{-1}\overline{\mathbf{F}}\sum_m \frac{\widetilde{F}_m}{\mu_m}\phi_m(x) \tag{12}$$

if the series converges. Here,

$$\overline{W} = \begin{bmatrix} \overline{w}_{ee} & \overline{w}_{ei} \\ \overline{w}_{ie} & \overline{w}_{ii} \end{bmatrix}, \qquad \overline{\mathbf{F}} = \begin{bmatrix} \overline{F}_e \\ \overline{F}_i \end{bmatrix},$$

$\widetilde{F}_m = \langle F, \phi_m \rangle$, *and* μ_m *is the eigenvalue of the integral operator* $[\mathcal{K}r](x) = \int_{\Omega} k(x,y)r(y)\,dy$ *associated with eigenfunction* ϕ_m.

Note that the first product in Eq. (12) (before the sum) is a vector (it has an excitatory and inhibitory component), but is not a function of x. On the other hand, the second term (the sum) is scalar, but depends on x. Hence, the solution is broken into its spatial and excitatory-inhibitory components.

If $\mathcal{K}$ has a nontrivial nullspace, then some of the eigenvalues μ_m will be zero. For Eq. (9) to be solvable in this case, $\widetilde{F}_m = \langle F, \phi_m \rangle$ must be zero for all such m and the corresponding terms in the series in Eq. (12) should be interpreted as zero.

The same separation of terms cannot be applied to solving Eq. (10), *i.e.*, Eq. (11) cannot be simplified in the same way when $\epsilon \neq 0$. However, the evaluation of Eq. (11) when $\epsilon \neq 0$ is simplified to some degree by noting that $\widetilde{W}_m = \overline{W}\mu_m$ and $\widetilde{\boldsymbol{F}}_m = \overline{\boldsymbol{F}}\widetilde{F}_m$ under the assumptions of Corollary 1 and therefore

$$\boldsymbol{r}(x) = \sum_m [\epsilon D - \overline{W}\mu_m]^{-1}\overline{\boldsymbol{F}}\widetilde{F}_m \phi_m(x). \tag{13}$$

Since all the examples we consider satisfy the assumptions of Corollary 1, we hereafter focus on Eqs. (12) and (13) in place of the more general Eq. (11). The analysis of specific networks with specific external input profiles can proceed as follows: The convergence of the series in Eq. (12) should be checked first. If it does not converge, then the network cannot maintain balance as $N \to \infty$ and Eq. (13) must be used at finite N instead. Even when Eq. (12) does converge, Eq. (13) can be used as an alternative to Eq. (12) for approximating firing rates in spiking network simulations. We next compare Eqs. (12) and (13) to results from large spiking network simulations.

4 Comparison of mean-field theory to spiking network simulations

We first consider the simulations discussed above and shown in Fig. 1. For this example, we can write $w_{ab}(x,y) = \overline{w}_{ab}k(x,y)$ where $\overline{w}_{ab} = 12\overline{p}_{ab}j_{ab}q_b$ and $k(x,y) = \min(x,y) - xy$, and we can write $F_a(x) = \overline{F}_a F(x)$ with $F(x) = \sin(\pi x)$. Therefore, we can apply Corollary 1. It is easily checked that the integral operator $[\mathcal{K}r](x) = \int_{\Omega} k(x,y)r(y)\,dy$ has eigenvalues and orthonormal eigenfunctions

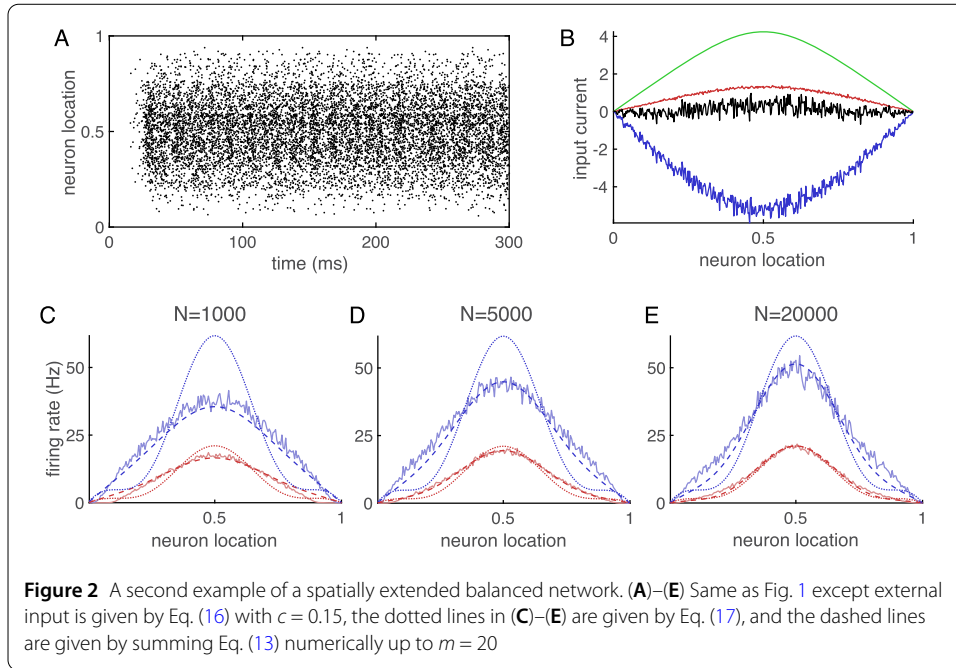$$\mu_m = \frac{1}{m^2\pi^2} \quad \text{and} \quad \phi_m(x) = \sqrt{2}\sin(m\pi x)$$

for $m \in \mathbb{N}$. Since $F(x) = \sin(\pi x)$ is itself an eigenfunction, its expansion in the eigenfunction basis has only one nonzero term $\widetilde{F}_1 = \langle F, \phi_a \rangle = 1/\sqrt{2}$ and $\widetilde{F}_m = 0$ for $m \neq 1$. Therefore, the series in Eq. (12) has only one nonzero term and reduces to

$$\boldsymbol{r}(x) = -\overline{W}^{-1}\overline{\boldsymbol{F}}\pi^2\sin(\pi x), \tag{14}$$

which represents the $N \to \infty$ limit of firing rates in the balanced state. The finite N correction in Eq. (13) similarly has only one term and is given by

$$\boldsymbol{r}(x) = \left[\pi^2\epsilon D - \overline{W}\right]^{-1}\overline{\boldsymbol{F}}\pi^2\sin(\pi x). \tag{15}$$

Figure 2 A second example of a spatially extended balanced network. (**A**)–(**E**) Same as Fig. 1 except external input is given by Eq. (16) with $c = 0.15$, the dotted lines in (**C**)–(**E**) are given by Eq. (17), and the dashed lines are given by summing Eq. (13) numerically up to $m = 20$

Comparing these expressions to firing rates computed from simulations shows that Eq. (15) provides a noticeably better approximation at $N = 1000$ (Fig. 1(C); dashed is closer than dotted to solid), but both equations agree very well at larger values of N (Fig. 1(D), (E); dashed and dotted are close to solid), as expected.

Figure 1(F), (G) shows individual neurons' firing rates versus their time-averaged input currents (dots) and compares them to the rectified linear fit used to estimate g_e and g_i (solid curves). Note that the rectified linear fit is not highly accurate, especially for excitatory neurons, because the true relationship between I and r is not a rectified linear function. However, this rough approximation is sufficient because g_e and g_i only affect Eq. (15) at order $\mathcal{O}(\epsilon)$.

We next consider a more interesting example in which the series in Eqs. (13) and (12) have infinitely many nonzero terms. In particular, consider the same recurrent network with external input
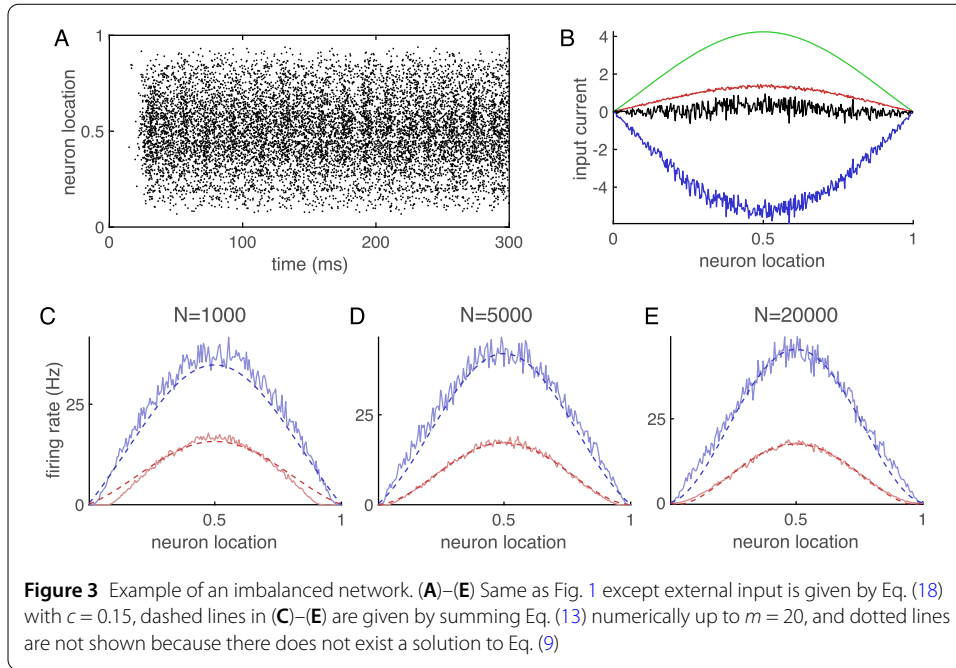
$$F(x) = c\sin^4(x) + (1 - c)\sin(\pi x). \tag{16}$$

Since we have not changed the network, μ_m and $\phi_m(x)$ are the same as above and the only term that changes from above is

$$\widetilde{F}_m = \frac{(1 - c)}{\sqrt{2}}\delta_{m,1} + \begin{cases} \frac{48c\sqrt{2}}{(64m - 20m^3 + m^5)\pi} & m \text{ odd}, \\ 0 & m \text{ even}, \end{cases}$$

where δ is the Kronecker delta. It is clear then that the series in Eq. (12) converges since $\widetilde{F}_m/\mu_m \sim m^{-3}$ as $m \to \infty$. Indeed, one can simplify the series in Eq. (12) to obtain

$$r(x) = -\overline{W}^{-1}\overline{F}\big((1 - c)\pi^2 \sin(\pi x)$$
$$+ 2c\pi^2\big[\cos(4\pi x) - \cos(2\pi x)\big]\big) \tag{17}$$

Figure 3 Example of an imbalanced network. (**A**)–(**E**) Same as Fig. 1 except external input is given by Eq. (18) with $c = 0.15$, dashed lines in (**C**)–(**E**) are given by summing Eq. (13) numerically up to $m = 20$, and dotted lines are not shown because there does not exist a solution to Eq. (9)

as the firing rates in the $N \to \infty$ limit in the balanced state. This is only a valid firing rate solution when $r(x) > 0$ for all $x \in [0, 1]$, which requires that c is sufficiently small. Specifically, c must be small enough that $(1 - c)\sin(\pi x) + 2c[\cos(4\pi x) - \cos(2\pi x)] \geq 0$ for all $x \in [0, 1]$. For the finite N correction in Eq. (13), we were unable to obtain a closed form expression for the series, but it can easily be summed numerically using the equations for μ_m and $\widetilde{F}_m$ derived above.

Network simulations show asynchronous-irregular spiking activity (Fig. 2(A)) and excitatory-inhibitory balance (Fig. 2(B)). Comparing our theoretical equations to firing rates from simulations shows that Eq. (13) is much more accurate than Eq. (17) even at larger values of N, but the convergence of the two equations to each other (and to results from simulations) is visible when comparing $N = 5000$ to $N = 20,000$ (Fig. 2(C)–(E)).

We finally consider the same recurrent network with external input

$$F(x) = c\sin^2(x) + (1 - c)\sin(\pi x). \tag{18}$$

In this case,

$$\widetilde{F}_m = \frac{(1 - c)}{\sqrt{2}}\delta_{m,1} + \begin{cases} \frac{4c\sqrt{2}}{\pi(4m - m^3)} & m \text{ odd,} \\ 0 & m \text{ even.} \end{cases}$$

Therefore, $F_m/\mu_m \sim m^{-1}$ as $m \to \infty$ and the series in Eq. (12) diverges. This implies that the network does not maintain excitatory-inhibitory balance as $N \to \infty$ because Eq. (9) does not admit a solution. Equation (10) still admits a solution for each N, given by Eq. (13), but this solution diverges as $N \to \infty$ ($\epsilon \to 0$).

Despite the break in balance as $N \to \infty$, network simulations still show asynchronous-irregular spiking activity (Fig. 3(A)) and approximate excitatory-inhibitory balance

(Fig. 3(B)) at finite N. Comparing Eq. (13) to firing rates from simulations shows that it is still somewhat accurate for multiple values of N (Fig. 3(C)–(E)).

Interestingly, the break in balance is not highly apparent in Fig. 3(B), as the currents still appear to be approximately balanced. Our mathematical analysis shows that the mean total input current cannot remain $\mathcal{O}(1)$ for all x as $N \to \infty$. However, we expect the divergence to grow like $\mathcal{O}(\sqrt{N})$ when balance is broken [12], so very large values of N could be necessary for imbalance to become visible.

5 Discussion

We extended the mean-field theory of firing rates in balanced networks to account for spatial connectivity structures in which connection probabilities depend on the spatial location of pre- and postsynaptic neurons without the translation invariance assumed in previous work. Any such network cannot maintain balance as $N \to \infty$ for every external input profile, and we derived conditions on the external input profile and connection probabilities required for balance to be possible. We also derived a finite N approximation to the firing rates that is applicable even when strict balance cannot be achieved as $N \to \infty$. We compared our theoretical results to large simulations of randomly connected integrate-and-fire neuron models. While the equations left some error at finite N, they captured the overall shape of firing rates at large values of N.

For balance to be realized, the firing rate profiles given by Eq. (12) need to be positive at all values of $x \in \Omega$. While we did not derive explicit conditions on this positivity, the solution in Eq. (12) can be checked for positivity and should only be interpreted as a valid solution when rates are positive. When Eq. (12) gives negative rates for some $x \in \Omega$, this would lead to mean firing approaching zero as $N \to \infty$ over some spatial locations. It is possible that the remaining neurons in the network that have nonzero firing rates could realize balance separately, forming a balanced sub-network. This possibility, which requires a nonlinear analysis, will be explored in future work.

We parameterized feedforward input as a time-constant, deterministic function $\boldsymbol{F}(x)$. In reality, cortical populations receive noisy, time-varying feedforward input from neurons in other cortical layers, cortical areas, and thalamus. These can be modeled more realistically by generating the spike trains of these presynaptic populations and assigning feedforward connectivity analogous to the recurrent connectivity considered here [15, 17, 22]. In the mean-field theory, this gives $\boldsymbol{F}(x) = [\mathcal{W}_F r_F](x)$ where $r_F(y)$ quantifies the spatial profile of firing rates in the presynaptic population, which might live on a different space $y \in \Omega_y$, and $\mathcal{W}_F$ is the mean-field connectivity kernel defined analogously to $\mathcal{W}$. In this case, Eq. (9) admits a solution, *i.e.*, balance can be achieved, whenever the range of $\mathcal{W}_F$ is contained within the range of $\mathcal{W}$. See [17] for a development of this idea for networks with translationally invariance connectivity profiles. This provides a way to test the spatial structure of multi-layered cortical circuits for the ability to maintain balance.

In addition, balance requires that the fixed point realized by Eq. (12) is stable. Stability of firing rate fixed points in networks of integrate-and-fire neurons can be very complicated and is outside the scope of this study. Stability for integro-differential equations of the form $\dot{\boldsymbol{r}} = -\boldsymbol{r} + g(\mathcal{W} + \boldsymbol{X})$ can be analyzed more easily, especially when g is assumed to be linear at the fixed point. This approach to stability analysis can provide a heuristic approximation to stability in spiking networks, but instabilities can arise in spiking networks that are not captured by the integro-differential equation due to the inherent resonance of

spiking neurons [29]. An in-depth analysis of stability and the dynamics that arise when the networks are destabilized would be an interesting direction for future studies.

We assumed that connectivity kernels are symmetric $w_{ab}(x, y) = w_{ab}(y, x)$ and that all four kernels share the same eigenfunctions. Extensions to this theory for asymmetric kernels and distinct eigenfunctions could be achieved using more general numerical or analytical approaches to solving integral equations that rely on singular value decompositions instead of eigenfunction decompositions (the two coincide under our assumptions) [21, 30, 31]. Such an extension would, for example, allow the analysis of visual cortical network models with pinwheel-shaped orientation hypercolumns. This could potentially be achieved analytically using methods to recast those networks in spherical coordinates [32] or numerically using real imaging of hypercolumn geometry and neural connectivity [18]. This could lead to studies that numerically or analytically probe which visual stimuli break excitatory-inhibitory balance. Since imbalanced networks generally lead to an amplification of firing rate responses [17], this would shed light on which visual stimuli patterns should be more salient under a balance network formalism.

In summary, our study extends the theory of balanced networks to more intricate spatial topologies, which opens the door to a number of additional lines of inquiry that could provide additional insights into the operation of cortical circuits.

Abbreviations
Not applicable.

Availability of data and materials
Matlab and C code for running all simulations and generating all figures is publicly available at
https://github.com/RobertRosenbaum/SpatialBalNetNonTransInvar/.

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Authors' contributions
CE and RR each contributed to all aspects of this project. All authors read and approved the final manuscript.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. van Vreeswijk C, Sompolinsky H. Chaotic balanced state in a model of cortical circuits. Neural Comput. 1998;10(6):1321–71.
2. van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science. 1996;274(5293):1724–6.
3. Shu Y, Hasenstaub A, McCormick DA. Turning on and off recurrent balanced cortical activity. Nature. 2003;423(6937):288–93.
4. Wehr M, Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature. 2003;426(6965):442–6.

5. 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):4535–45.
6. Okun M, Lampl I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat Neurosci. 2008;11(5):535–7.
7. Dorrn AL, Yuan K, Barker AJ, Schreiner CE, Froemke RC. Developmental sensory experience balances cortical excitation and inhibition. Nature. 2010;465(7300):932–6.
8. Sun YJ, Wu GK, Liu B-H, Li P, Zhou M, Xiao Z, Tao HW, Zhang LI. Fine-tuning of pre-balanced excitation and inhibition during auditory cortical development. Nature. 2010;465(7300):927–31.
9. Zhou M, Liang F, Xiong XR, Li L, Li H, Xiao Z, Tao HW, Zhang LI. Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. Nat Neurosci. 2014;17(6):841–50.
10. Petersen PC, Vestergaard M, Jensen KHR, Berg RW. Premotor spinal network with balanced excitation and inhibition during motor patterns has high resilience to structural division. J Neurosci. 2014;34(8):2774–84.
11. Monteforte M, Wolf F. Dynamic flux tubes form reservoirs of stability in neuronal circuits. Phys Rev X. 2012;2(4):041007.
12. Rosenbaum R, Doiron B. Balanced networks of spiking neurons with spatially dependent recurrent connections. Phys Rev X. 2014;4(2):021039.
13. Landau ID, Egger R, Dercksen VJ, Oberlaender M, Sompolinsky H. The impact of structural heterogeneity on excitation-inhibition balance in cortical networks. Neuron. 2016;92(5):1106–21.
14. Pyle R, Rosenbaum R. Highly connected neurons spike less frequently in balanced networks. Phys Rev E. 2016;93(4):040302.
15. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B. The spatial structure of correlated neuronal variability. Nat Neurosci. 2017;20(1):107.
16. Darshan R, van Vreeswijk C, Hansel D. Strength of correlations in strongly recurrent neuronal networks. Phys Rev X. 2018;8(3):031072.
17. Ebsch C, Rosenbaum R. Imbalanced amplification: a mechanism of amplification and suppression from local imbalance of excitation and inhibition in cortical circuits. PLoS Comput Biol. 2018;14(3):1006048.
18. 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.
19. Lund JS, Angelucci A, Bressloff PC. Anatomical substrates for functional columns in macaque monkey primary visual cortex. Cereb Cortex. 2003;13(1):15–24.
20. Large AM, Vogler NW, Canto-Bustos M, Friason FK, Schick P, Oswald A-MM. Differential inhibition of pyramidal cells and inhibitory interneurons along the rostrocaudal axis of anterior piriform cortex. Proc Natl Acad Sci. 2018;115(34):8067–76.
21. Tricomi FG. Integral equations. New York: Interscience; 1957.
22. Renart A, de La Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris K. The asynchronous state in cortical circuits. Science. 2010;327(5965):587–90.
23. Barral J, Reyes AD. Synaptic scaling rule preserves excitatory–inhibitory balance and salient neuronal network dynamics. Nat Neurosci. 2016;19(12):1690.
24. Ermentrout B. Neural networks as spatio-temporal pattern-forming systems. Rep Prog Phys. 1998;61(4):353.
25. Bressloff PC. Spatiotemporal dynamics of continuum neural fields. J Phys A. 2012;45(3):033001.
26. Brunel N, Hakim V. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. Neural Comput. 1999;11(7):1621–71.
27. Brunel N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci. 2000;8(3):183–208.
28. Lindner B. Effects of noise in excitable systems. Phys Rep. 2004;392(6):321–424.
29. Pyle R, Rosenbaum R. Spatiotemporal dynamics and reliable computations in recurrent spiking neural networks. Phys Rev Lett. 2017;118(1):018103.
30. Hansen P. Numerical tools for analysis and solution of Fredholm integral equations of the first kind. Inverse Probl. 1992;8(6):849.
31. Hansen PC. Discrete inverse problems: insight and algorithms. vol. 7. Philadelphia: Society for Industrial and Applied Mathematics; 2010.
32. Bressloff PC, Cowan JD. A spherical model for orientation and spatial–frequency tuning in a cortical hypercolumn. Philos Trans R Soc Lond B, Biol Sci. 2002;358(1438):1643–67.