

Targeting diverse operational regimes in recurrent spiking networks

Pierre Ekelmans^{1,2}, Nataliya Kravnyukova^{1,3}, and Tatjana Tchumatchenko^{1,3,4}

¹ *Max Planck Institute for Brain Research Theory of neural dynamics group*

² *Frankfurt Institute for Advanced Studies*

³ *University of Bonn Medical Center Institute of experimental epileptology and cognition research*

⁴ *University of Mainz Medical Center Institute of physiological chemistry*

Neural computations emerge from recurrent neural circuits that comprise hundreds to a few thousand neurons. Continuous progress in connectomics, electrophysiology, and calcium imaging require tractable spiking network models that can consistently incorporate new information about the network structure and reproduce the recorded neural activity features. However, it is challenging to predict which spiking network connectivity configurations and neural properties can generate fundamental operational states and specific experimentally reported nonlinear cortical computations. Theoretical descriptions for the computational state of cortical spiking circuits are diverse, including the balanced state where excitatory and inhibitory inputs balance almost perfectly or the inhibition stabilized state (ISN) where the excitatory part of the circuit is unstable. It remains an open question whether these states can co-exist with experimentally reported nonlinear computations and whether they can be recovered in biologically realistic implementations of spiking networks. Here, we show how to identify spiking network connectivity patterns underlying diverse nonlinear computations such as XOR, bistability, inhibitory stabilization, supersaturation, and persistent activity. We established a mapping between the stabilized supralinear network (SSN) and spiking activity which allowed us to pinpoint the location in parameter space where these activity regimes occur. Notably, we found that biologically-sized spiking networks can have irregular asynchronous activity that does not require strong excitation-inhibition balance or large feedforward input and we showed that the dynamic firing rate trajectories in spiking networks can be precisely targeted without error-driven training algorithms.

I. INTRODUCTION

Layered or columnar neuronal structures consisting of hundreds to thousands of neurons constitute local computational blocks in the mammalian cortex. Each computational block has its particular size and connectivity rules, which determine its dynamics and computational repertoire. Therefore, understanding the computational regimes of recurrent networks with different sizes ranging from hundreds to millions of neurons and permitting diverse connectivity patterns is essential to explain the emergence of cognitive functions and behavior. It is currently a challenge to quantitatively predict the activity of medium-sized spiking neural networks. The most powerful mathematical theories now operate at two opposite scales: they either model a small number of neurons that generate a particular activity pattern of interest [1], or they operate in the limit of large or infinitely large networks [2]. Few theories can make mathematically tractable, quantitative, and experimentally relevant predictions for the diverse, intermediate sizes of spiking networks reported for local cortical circuits [3–5].

Here, we study the activity regimes of spiking networks whose sizes range from a few hundred to thousands of neurons. Many parameters describing spiking neurons and their intracortical connections have recently been measured across cortical cell types [6], and detailed numerical network simulations have been put forward [7]. However, it is challenging to interpret spiking network simulations because network dynamics depend strongly on multidimensional parameter settings, while exper-

imentally reported parameters can vary within broad ranges [8–12]. On the opposite side of the complexity spectrum are population rate networks [13–15] which describe the average activity of neurons in each population and can relate a specific activity regime to a connectivity configuration. However, it is still unclear how to translate computational results obtained with rate models to biologically plausible spiking networks.

One of the most popular models for large or infinitely large networks is the balanced state framework [16]. Its biological correlate is the experimentally reported strong balance between excitatory and inhibitory synaptic currents [17] and asynchronous irregular spiking activity [18]. However, the computational hallmarks of the balanced network limit, including response linearity and strong feedforward connections, are not consistent with the diversity of experimentally reported non-linear responses across cortical areas [19] and reports of weak feedforward inputs [20]. Therefore, experimental observations deviating from the balanced state predictions have been often addressed by keeping the core excitatory/inhibitory (E/I) balance framework while adding specific synaptic plasticity rules [21], more complex synaptic strength distributions [22], or a semi-balance condition which allows for excess inhibition [23]. Furthermore, the existence and stability of the balanced state limit come with strict conditions on the feedforward and recurrent connection strengths (see [16, 24], Appendix), which exclude many network connectivity configurations. Finally, the balanced limit is based on the assumption that networks are infinitely large and ignores finite-size

effects.

Firing rate models of neural sub-populations with a nonlinear activation function [13, 25, 26] are an alternative to the balanced state model. These models often use either a threshold linear function [25] or an S-shaped sigmoidal function [13, 27] to describe the activation function (F-I curve) of neurons. The advantage of the threshold linear network model is that it can be solved analytically; however, it fails to reproduce nonlinear responses often observed in experiments. On the other hand, network models with the sigmoidal transfer functions support nonlinear responses; however, systematic mathematical analysis for the activity states of these networks is currently in short supply. The stabilized supralinear network model (SSN) [26] uses a supralinear power law as a transfer function. The advantage of the SSN model is that its activity states can be characterized analytically [15, 28] and it can reproduce a variety of nonlinear cortical responses in the realistic range of firing rates observed *in vivo* [26].

Could the SSN model predict and quantify the activity regimes in medium-sized spiking networks for all connectivity configurations? Here, we argue that the SSN model can be used to predict diverse nonlinear responses such as supersaturation, bistable activity, and inhibition stabilized regimes in spiking networks. We propose a mapping between the parameter space of the SSN model and the more complex parameter space of the leaky-integrate-and-fire (LIF) network of spiking neurons. The advantage of this approach is that it results in a mathematically tractable model which can be manipulated analytically. Thanks to this, we can invert the model equations and design the network's input to target a desired activity trajectory or operational regime of interest via closed-form equations without network training. This approach can be used to generate specific nonlinear functions (eg: XOR gate).

In relation to the balanced state, we show that for biologically realistic connectivity configurations, even very large networks (size $N \approx 10^5$) can deviate substantially from the balanced state predictions despite high E-I correlations. Likewise, networks lacking precise E-I input balance can produce asynchronous irregular firing patterns resembling balanced network activity. Overall, we find that medium-sized spiking networks can have a complex, nonlinear behavior that is accurately described with the SSN framework and can be far from the predictions of the balanced state limit.

Examining the possible activity regimes of networks with biologically plausible synaptic strengths, we could delineate different computational regimes such as supersaturation and bistability. We found that only a specific range of connectivity and input choices allow the networks to be in the inhibition stabilized state [25]. The inhibition stabilized state has been suggested as an operational state of the cortex [25, 29] and is linked to a decrease of the inhibitory firing rates following the increase in inhibitory input. Here, we delineate the parameter

regimes in the medium-sized networks where inhibitory stabilization and a set of other experimentally reported nonlinear computations co-exist and parameter regions where they are mutually exclusive.

II. RESULTS

Our goal is to understand how neural circuits comprising a few thousand neurons organize their spiking activity. We want to predict whether specific nonlinear computations can occur in these networks and pinpoint their location in the multidimensional parameter space spanned by recurrent connectivity and input weights. We choose the size of the networks to be 4000 neurons, which is biologically plausible for local cortical circuits (Appendix). Similarly, we restrict our analysis to the *in vivo* reported range of activity of 0-10 Hz [30–37]. We choose the strength and probability of synaptic connections to be within the same order of magnitude as the values reported by the database of the Allen Institute for the visual cortex area V1 in mice [6] (Appendix). To model cortical activity, we use the leaky-integrate-and-fire (LIF) model (see Methods), which represents an accurate description of cortical neurons both *in vivo* and *in vitro* [38]. To predict spiking network activity regimes, we map the mean activity of spiking networks to a rate-based 2D SSN model.

A. Approximating spiking network activity with the SSN model

Our starting point is the observation that a power-law function can accurately describe the F-I curve of a single LIF neuron across different membrane time constants τ and input noise values σ (see Fig. 1B,C). The power-law approximation is consistent with the Ricciardi exact solution of the LIF neuron Eq. 9 [39], which we denote by Φ (Fig. 1B). Throughout this study, we are interested in the firing rate range from zero to approximately 10 Hz which has been reported *in vivo* across many brain areas [30–37]. To this end, we fit the low firing rate regime (0-10 Hz) of the F-I curve of a LIF neuron given by Φ using the threshold power-law function of the form

$$\nu = a(\mu - b)_+^n. \quad (1)$$

Where ν is the firing rate, μ is the input to the neuron and $(x)_+ = \max\{x, 0\}$. The constants a , b , and n characterize the power-law approximation, with a scaling pre-factor a , an input threshold b upon which the neuron starts firing, and an exponent n . The power-law exponent n in our approximation varies between 2 and 4, which is consistent with the biologically reported range [35, 40].

We connect the individual LIF neurons into a recurrent network of excitatory (E) and inhibitory (I) neurons (Fig. 1A). We assume that E and I neurons differ in their membrane time constants ($\tau_E=20$ ms, $\tau_I=10$ ms,

black crosses in Fig. 1C) consistently with experimental reports [6]. We note that the input to a neuron in a recurrent network, which is a superposition of postsynaptic potentials (PSPs), is equivalent to an Ornstein Uhlenbeck process or white noise if the number of incoming PSPs is sufficiently large and the activity is irregular [41, 42].

To describe the activity of the E and I populations, we use the power-law approximation of the single-neuron transfer function (Eq. 1). In the mean field approximation, the average firing rate of each population is given by a system of equations equivalent to the SSN [26]

$$\begin{aligned}\tau_{PE} \frac{d\nu_E}{dt} &= -\nu_E + a_E(\mu_E - b_E)_+^{n_E} \\ \tau_{PI} \frac{d\nu_I}{dt} &= -\nu_I + a_I(\mu_I - b_I)_+^{n_I}.\end{aligned}\quad (2)$$

Where ν_X , $X \in \{E, I\}$ are the firing rates of the two populations, the parameters a_X , b_X , and n_X are given by the F-I curve fit of E and I neurons (black crosses in Fig. 1C), and μ_X represent the sum of recurrent and feedforward inputs to each population. The population time constants τ_{PX} characterize how fast the firing rate of each population evolves. *In vivo* cortical networks have been shown to respond to sudden stimulation with a transient (or *onset*) response that has a time scale of approximately 20 ms [43]. Furthermore, *in vivo* recordings of multiple cortical areas have reported autocorrelation timescales of the order of hundreds ms [44–46] suggesting that slow timescales that are larger than the neuronal or synaptic variables are relevant for the analysis of biological circuits. Therefore, in the following we will consider firing rate dynamics that evolve on the time scales of a few tens to a hundred milliseconds which is slower than the intrinsic time scales of the network. Based on these time scales, the firing rate dynamics of the recurrent network can be understood using the equilibrium state and the external drive of a network. The equilibrium states of the SSN equation are described by

$$\begin{aligned}\nu_E &= a_E(J_{EE}\nu_E - J_{EI}\nu_I + \mu_{\text{ext}} - b_E)_+^{n_E} \\ \nu_I &= a_I(J_{IE}\nu_E - J_{II}\nu_I + r\mu_{\text{ext}} - b_I)_+^{n_I}.\end{aligned}\quad (3)$$

Here r is the ratio of the external inputs to the I and E populations $r = \mu_{\text{extI}}/\mu_{\text{extE}}$, which allows for the simplified notation: $\mu_{\text{ext}} = \mu_{\text{extE}}$ and $r\mu_{\text{ext}} = \mu_{\text{extI}}$. The population-wise connection strengths J_{XY} characterize the recurrent connections from population Y to population X , whereby $X, Y \in \{E, I\}$.

Previous work identified the constraints on connectivity configurations in the SSN model that underlie such nonlinear activity responses as supersaturation [15], the paradoxical effect [47, 48], bistability, and persistent activity [28]. We show that the parameters of LIF spiking networks can be mapped to the SSN such that the same activity types emerge in the spiking network, according to the observations made with the SSN. In the following sections, we discuss each activity type and its corresponding connectivity regime in the SSN, as well as in LIF spiking networks.

B. Supersaturation - Firing rates can decline for growing input

Firing rates of neurons *in vivo* can show a range of nonlinear behaviors as a function of stimulus strength [49]. In particular, the activity level of sensory neurons may decrease after stimulus onset, and a substantial number of pyramidal V1 neurons in mice show reduced firing in response to enhanced stimulus contrast [50]. At the same time, the average activity of thalamic neurons in mice - primarily targeting V1 neurons - is an increasing function of the stimulus contrast [51]. Therefore, it appears that E neurons can be suppressed despite the increase in external input. This phenomenon is generally referred to as supersaturation [15].

First, we studied E firing response to growing inputs and aimed to delineate parameter regimes where a decreasing population response can be observed. We found that supersaturation ($\frac{d\nu_E}{d\mu_{\text{ext}}} < 0$) can be observed in a large class of connectivity and input weights configurations within the spiking networks that can be predicted by the inequality derived for the SSN model in [15, 52]

$$r > \frac{J_{II}}{J_{EI}}.\quad (4)$$

Interestingly, only three network parameters determine the SSN network's ability to be in a supersaturating activity regime (Eq. 4). For a network to be supersaturating, the ratio of external inputs r has to exceed the ratio of recurrent inhibition J_{II}/J_{EI} . As a result, the remaining two connectivity parameters (the recurrent excitation J_{IE} and J_{EE}) cannot control the existence of supersaturating activity in the SSN. The exact point at which a network satisfying Eq. 4 becomes supersaturating does, however, depend on all network parameters as it occurs when the inhibitory firing rate exceeds a specific threshold value (Appendix, Eq. B5).

To understand if the connectivity condition derived in the SSN model (Eq. B5) leads to a quantitative description of supersaturation in spiking networks, we generated LIF network parameters fulfilling the supersaturating condition using the SSN-LIF mapping framework in Eq. 1 (Fig. 2A). We found that the activity in a LIF spiking network and its self-consistency approximation Φ_{sc} (Eq. 12) align robustly with the activity of the SSN model (Fig. 2A).

Recent work [27] compared the responses of LIF and SSN models, pointing out that the peak E activity in supersaturating spiking networks is small, and in particular, it is smaller than the SSN peak. As shown in Fig. 2A, the peak firing rates obtained with the two methods are in agreement. Furthermore, we show that it is possible to control the height of the E firing rate peak in both networks such that it can be made arbitrarily high (Fig. 2B, Appendix). Specifically, we show that the peak of E activity can be controlled by modifying the ratio of the external inputs r , and the connectivity parameters J_{EI} and J_{IE}^{-1} by the same factor. This manipulation derived from

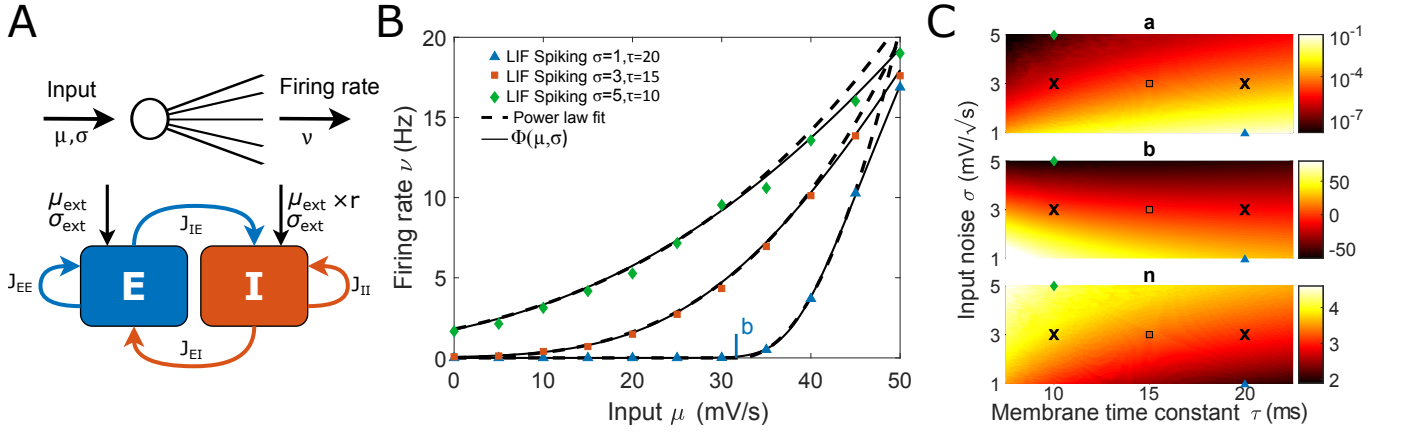


FIG. 1. **Spiking neurons can be quantitatively described by a supralinear power law for low activity.** (A, top) Schematic representation of the F-I transfer function of a neuron. (A, bottom) Architecture of the recurrent Excitatory-Inhibitory network. (B) Neuronal firing rate as a function of input for different input noise σ and membrane time constant τ . The firing rate response of simulated LIF neurons (colored triangles) is in line with the Ricciardi solution Φ (solid line, Eq. 9). The power-law approximation (Eq. 1) accurately aligns with the LIF simulation and Ricciardi solution for low firing rates. Note that the power-law fit is only applied in the range of $\nu < 10$ Hz, and diverges beyond this range. The vertical mark denotes the power-law parameter b for one of the curves (b is negative for the other two curves). (C) The power-law parameters depend on input noise σ and the membrane time constant τ . The two crosses indicate the parameter regimes we use for the excitatory ($\tau_E = 20$ ms) and inhibitory ($\tau_I = 10$ ms) neurons in recurrent networks. Other symbols indicate the parameters τ and σ used in B. The fit is obtained with the least squares method. Power-law parameters are listed in Table III.

the SSN analysis (Appendix) leads to the same effect in the spiking networks (Fig. 2B).

To determine how close the network operates to E-I balance, we measure the Pearson correlation of the time series of recurrent excitatory and inhibitory inputs to each neuron $\rho_{E,I}$. The E-I correlation measured at the peak E firing rate is close to zero, demonstrating that the network operates far from E-I balance (Fig. 2C). Since the coefficient of variation of the interspike intervals at the peak E firing rate is close to 1 ($CV_{ISI} \approx 1$), the firing appears to be irregular, asynchronous, and compatible with a Poisson spiking process (Fig. 2D). Importantly, the supersaturation regime occupies the biologically plausible activity range of 0-10 Hz in spiking networks [30–37], and the amplitude of the synaptic connection strengths, as well as the size of the network ($N = 4000$), are both in line with biological estimates of functional cortical network size [6, 53, 54] (Appendix). We note that the supersaturation condition is incompatible with the existence of a balanced state solution, as defined in Eq. B1 as it would lead to negative firing rates.

Knowing how the 2D firing rates emerge from recurrent and feedforward connectivity in the SSN allows us to invert this relation and select external inputs such that they lead to the desired E and I activity trajectory in the spiking network. This is illustrated in Fig. 2E where we targeted a complex 2D trajectory. We obtained the feedforward inputs that result in the desired dynamics $\nu_E(t)$

and $\nu_I(t)$, by inverting Eq. 3:

$$\begin{aligned} \mu_{extE}(t) &= \left(\frac{\nu_E(t)}{a_E} \right)^{1/n_E} + J_{EI}\nu_I(t) - J_{EE}\nu_E(t) \\ \mu_{extI}(t) &= \left(\frac{\nu_I(t)}{a_I} \right)^{1/n_I} + J_{II}\nu_I(t) - J_{IE}\nu_E(t). \end{aligned} \quad (5)$$

These dynamic feedforward inputs $\mu_{extE}(t)$ and $\mu_{extI}(t)$ are shown in Fig. 2F, bottom and the fidelity of the targeting is illustrated in Fig. 2F, top. Notably, the timescale of the autocorrelation function of neuronal activity (as defined in [55]) is around 300 ms, which is in line with recorded cortical activity [44–46]. These results indicate that complex dynamic trajectories evolving on biologically realistic timescales can be accurately captured by the SSN steady states Eq. 3.

Let us note that while we used here dynamic feedforward inputs to move along the activity trajectory, it is equally possible to dynamically modify the connectivity to obtain the same 2D trajectory in activity space. In this scenario, synaptic plasticity is recruited to obtain a user-defined output. This can be done by setting the plastic connections J as dynamic while the external inputs are constant.

$$\begin{aligned} J_{EE}(t) &= \frac{\left(\frac{\nu_E(t)}{a_E} \right)^{\frac{1}{n_E}} + J_{EI}\nu_I(t) - \mu_{ext} + b_E}{\nu_E(t)} \\ J_{IE}(t) &= \frac{\left(\frac{\nu_I(t)}{a_I} \right)^{\frac{1}{n_I}} + J_{II}\nu_I(t) - r\mu_{ext} + b_I}{\nu_E(t)}. \end{aligned} \quad (6)$$

Overall, we show that the mapping between SSN and spiking networks makes it possible to construct inputs or synaptic weights in a spiking neural circuit such that its activity follows a user-defined complex target dynamical trajectory..

In balanced networks, the implementation of logical gates is a complex task due to the linearity of the transfer function [23]. Therefore, we asked whether the nonlinear regimes of spiking networks can be used to perform specific logical operations. Here, we show that it is possible to combine feedforward and recurrent inputs in a way that makes the circuits perform the nonlinear XOR operation, which is one of the key computing components of logical circuits, while being challenging to implement in a neural network [56]. We show in Fig. 2G how a supersaturating network can execute the XOR operation from two input signals. The E activity is maximal if the input $X_1 + X_2$ corresponds to the peak input in the SSN supersaturating regime. The E activity is unstimulated if both inputs X_1 and X_2 are low and silenced if they are both high. This shows that the nonlinearity of medium-sized spiking networks can be exploited to carry out fundamental logical operations.

Next, we explore if the experimentally reported connectivity parameters in mouse V1 by the Allen Institute [6] are consistent with supersaturation (Appendix, Table II). We use these parameters as a reference point for the biologically plausible range of connection strengths. Interestingly, a circuit with these connectivity parameters does not have a balanced state solution for the equal external input ratio $r = 1$ ($\mu_{\text{extE}} = \mu_{\text{extI}}$) and requires $r < 0.9$ to fulfill the balanced state requirement (Eq. B1). The network can be supersaturating for values of r larger than 0.9. Remarkably, for $r = 1$, the E activity does not decrease in the low input range but saturates instead (Fig. 3A), the network only becomes supersaturating for $\mu_{\text{ext}} > 150$ mV/s (Fig. S1B). For larger values of r the activity decreases and becomes silent for inputs close to 100 mV/s (see inset Fig. 3A, $r = 1.5$).

C. Inhibitory stabilization and its presence for reconstructed synaptic weights

Inhibitory stabilization is a network state in which the recurrent excitation feedback loop is strong and intrinsically unstable but can be stabilized by the recurrent inhibition [25, 57]. The paradoxical effect is a feature of the ISN [25], in which the I activity decreases as the input to the I population is increased ($\frac{d\nu_I}{d\mu_{\text{extI}}} < 0$). Recent studies using optogenetic stimulation of inhibitory neurons confirmed the paradoxical effect in mouse visual, somatosensory, and motor cortices [58] suggesting that the ISN is a ubiquitous property present across cortical networks. A recent review presented further experimental evidence and techniques used to study the inhibition-stabilized dynamics and discussed the ISN consequences for cortical computation [29]. In the SSN model [47, 48],

a network is inhibition-stabilized if it fulfills the condition

$$\nu_E > (a_E n_E^{n_E} J_{EE}^{n_E})^{-\frac{1}{n_E-1}}. \quad (7)$$

We note that in networks with a threshold linear transfer function, the analogous ISN condition only requires a strong recurrent coupling $J_{EE} > 1$ and does not impose any constraints on the E firing rate level or the transfer function parameters [25, 29]. However, large enough J_{EE} does not always guarantee that a recurrent neural network with a nonlinear transfer function is in the ISN regime. Increasing J_{EE} can also lead to instability, as the excitatory feedback loop can strengthen to a point where it escapes stabilization from recurrent inhibition. In the extreme case, it is even possible to build a network that can never enter the ISN regime regardless of the value of J_{EE} , as E activity never reaches the level where it can be stabilized by inhibition (Appendix).

Next, we investigated whether this condition (Eq. 7) can predict the existence of the ISN in spiking networks of LIF neurons. Interestingly, we found that the ISN condition cannot be met for the connectivity strengths reported for mouse V1 from the Allen Atlas [6] if the E/I input ratio r is equal to 1. This is due to the fact that the required E firing rate ($\nu_E > 27$ Hz) is higher than the maximal possible stable E firing rate in the network (Fig. 3A). For very low values of r (around $r = 0.1$), an ISN state can only be reached by exposing the network to very high external inputs (around $\mu_{\text{ext}} = 1000$ mV/s) (Fig. S1B). We will choose a network which operates outside these cases since the corresponding firing rates are far beyond the 0-10 Hz firing rate range we consider in this study. Therefore, to meet the ISN condition, we modified one of the connectivity strengths. Specifically, we increased the connectivity parameter J_{EE} , which is supported by the study by [10] who report larger J_{EE} than the Allen Atlas [6]. We chose J_{EE} such that the network is in the ISN state for E firing rates larger than 1.5 Hz and kept all other connectivity strengths as reported by the Allen Atlas ([6], see Table II). Fig. 3B shows that the resulting network exhibits the paradoxical effect and is therefore in the ISN regime.

D. Bistability and persistent activity

One of the most prominent experimentally recorded neural activity features *in vivo* is the network ability to switch between higher and lower firing levels. One example is spontaneously alternating intervals of tonic firing and silence observed across different cortical areas [34]. Another example is the sustained firing rate in the prefrontal cortex after stimulus withdrawal during decision-making tasks which is hypothesized to represent short-term memory [14, 59]. The coexistence of multiple network states can be explained theoretically by bistability, where the system has two stable states for the same level of input. If multiple stable states co-exist in a network model, a sufficiently large perturbation can drive

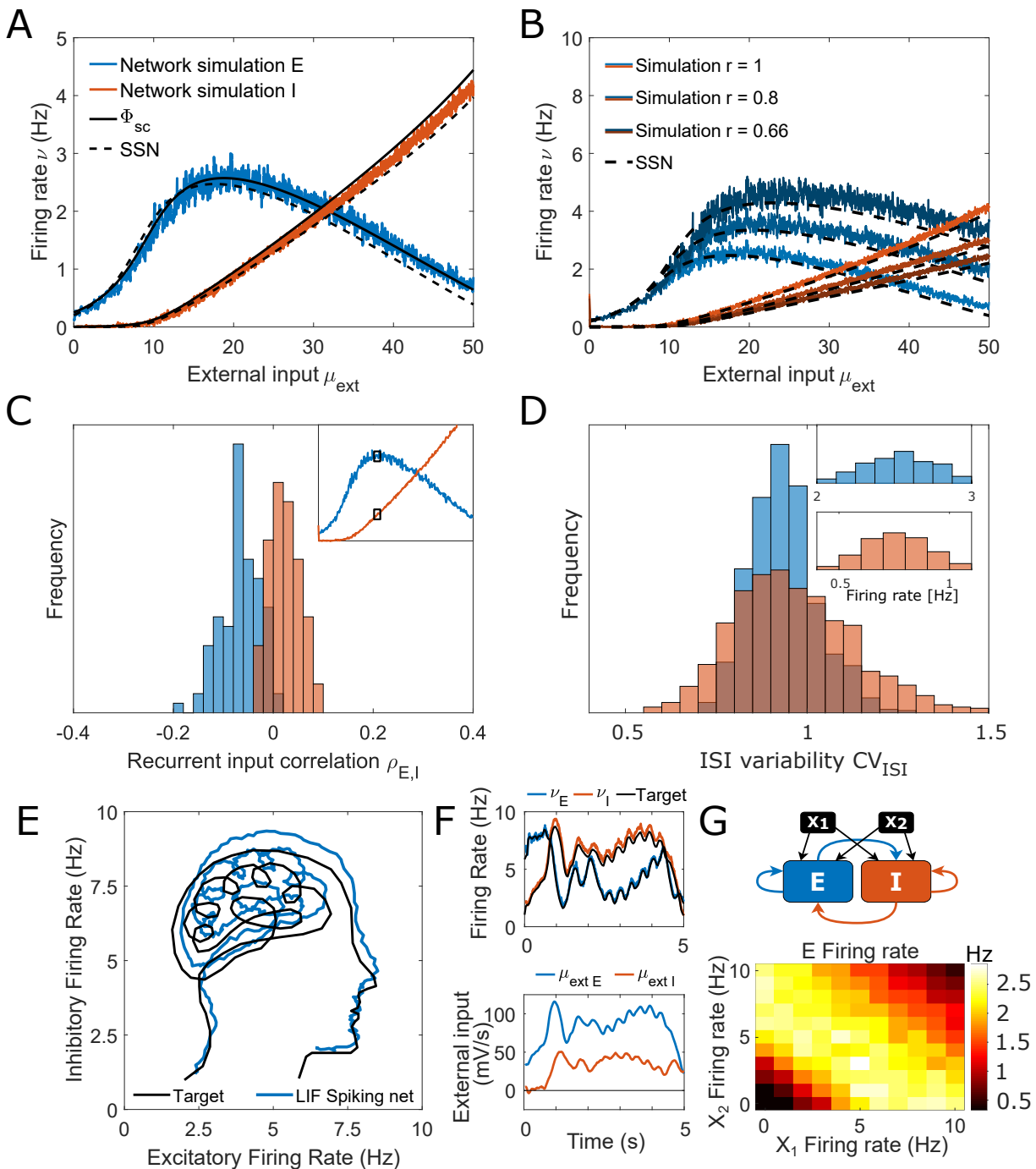


FIG. 2. SSN-predicted supersaturation can be observed in spiking network simulations. The existence of the supersaturating response can be predicted by the SSN framework using Eq. B5. (A) E (blue) and I (red) firing rates as a function of external input. LIF spiking activity is in line with the self-consistency solution (Φ_{sc} in Eq. 12, solid line) and is also accurately captured by the SSN solution (dashed line). (B) The peak of E activity can be tuned to any desired level by tuning the E/I ratio of external inputs (r) along with the connectivity weights J_{IE} and J_{EI} . (C) Histogram of the E/I input correlation ($\rho_{E,I}$) onto E (blue) and I (red) neurons computed at the peak of E activity (see inset). The weak correlation suggests the network is operating far from E/I balance. (D) The spiking activity of both E and I neurons is irregular and compatible with a Poisson process (CV_{ISI} close to 1). The inset shows the distribution of firing rates of individual neurons. (E) Spiking networks can follow a user-defined target dynamical trajectory. The black line shows the target trajectory we aim to replicate with the network. The blue line shows the trajectory of the spiking network in the E-I activity phase space. (F) Same simulation as in panel E, The time course of the E and I firing rates in the LIF network (top) follows the target trajectory and results from designed dynamical inputs (bottom). (G) Supersaturating networks can perform the XOR task. top: Layout of the network used to perform the XOR task, where the E-I network is supersaturating (Eq. 4). Bottom: The LIF E population activity performs XOR logical operation of the two inputs X_1 and X_2 . The feedforward weights are $J_{EX_1} = J_{IX_1} = J_{EX_2} = J_{IX_2} = 2.5$ mV. The spiking network parameters can be found in Table II.

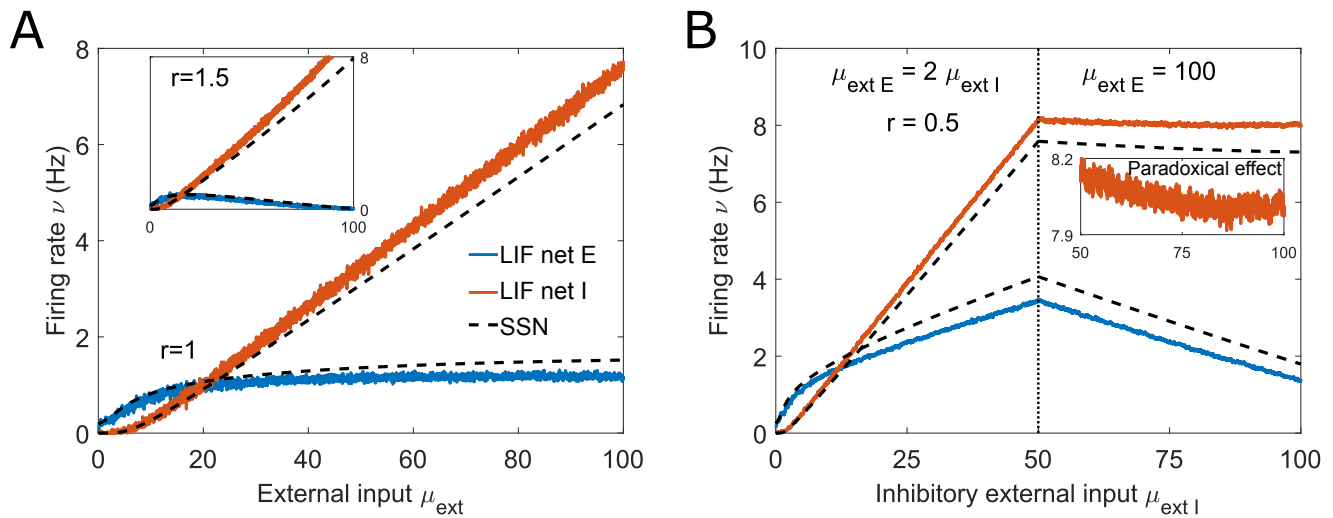


FIG. 3. The experimentally reported network parameters can generate supersaturation and be adapted to enter the inhibition-stabilized state. (A) Firing rates of the E and I populations, blue and red lines, respectively, as a function of external input using the parameters reported by the Allen institute [6], see Table II. The dashed lines show the SSN solution. In this network, the E activity saturates for inputs larger than 20 mV/s. If the external input to the I population is larger than that to the E population, as shown in the inset with $r = 1.5$, E firing rate declines for growing input. (B) This spiking network exhibits inhibitory stabilization matching the predictions of the SSN (dashed line). The connection strength J_{EE} is higher than in panel A (Table II). The ISN state is exposed by the paradoxical effect which occurs when the I firing rate decreases for increasing $\mu_{\text{ext I}}$. First, the inputs to both populations grow to drive the network in a state where the E subnetwork is unstable: $\nu_E > 1.5$ Hz, Eq. 7 (vertical dotted line). Once $\mu_{\text{ext I}}$ reaches 50 mV/s, only the input to I increases (from 50 to 100 mV/s), while $\mu_{\text{ext E}}$ remains at 100 mV/s. This results in a decrease in the firing rates of both populations, as predicted by the SSN (dashed lines). The inset shows a close up of the I activity to illustrate the paradoxical effect and shows that the paradoxical effect wanes as the E firing rate approaches the ISN threshold ($\nu_E \approx 1.5$ Hz).

network activity away from its current state towards another attractor. In the situation where a bistable network can sustain its high activity level in the absence of feedforward input, it has persistent activity. Here, we asked whether the SSN model can predict the connectivity regime supporting bistability in spiking circuits.

Bistability and persistent activity can be obtained in the SSN model [28] without the need for synaptic plasticity [21] or complex synaptic weight distributions [22]. Unlike supersaturation and inhibition stabilization, bistability cannot be delimited by a simple tractable condition on network parameters (Appendix). However, we can use the conditions presented in [28], as a starting point to guide our search for bistability in biologically realistic spiking networks, even though they are derived under restrictive assumptions on the a , b , and n parameters. We show an example of a biologically realistic bistable network in Fig. 4.

Both the self-consistency solution and LIF network simulation confirm the SSN-predicted bistability Fig. 4A: the network can sustain either low activity or high activity for external inputs in the 2-4 mV/s range. Although the SSN and Φ_{sc} are deterministic, the spiking network simulation is not. Due to the stochastic nature of the neuronal activity, fluctuations in firing can cause spontaneous transitions between steady states (shown in Fig. 4A, inset). We note that the spontaneous transi-

tions between the up and down states have not been reported in the bistable balanced networks with short-term plasticity [21] because spontaneous fluctuations in average firing and the probability for spontaneous transitions decrease with network size.

We find that a higher excitatory membrane time constant broadens the window of bistability (Fig. 4B), making bistability more robust to spontaneous fluctuations and easier to locate in phase space. As τ_E increases, the bistability window shifts towards lower feedforward input. When the bistability window intersects the vertical $\mu_{\text{ext}} = 0$ axis, the network has a persistent activity state in the absence of feedforward input ($\tau_E = 22$ ms in Fig. 4B). Here again, the E-I balance is weak, as shown by the correlation in recurrent inputs (Fig. 4C), and the spiking activity is Poisson-like, as shown by the coefficient of variation of the ISIs (Fig. 4C).

Finally, the nonlinear transformation performed by spiking networks can be functionally relevant for information processing. Logical operations such as the AND operation can be implemented without the need to recruit synaptic plasticity, thanks to the sharp transition between the two stable states. If the transition from the low to the high activity level requires a strong input, so that two signals X_1 and X_2 need to be present to elicit the transition, the network can execute the AND operation. Moreover, the bistability of a neural network can

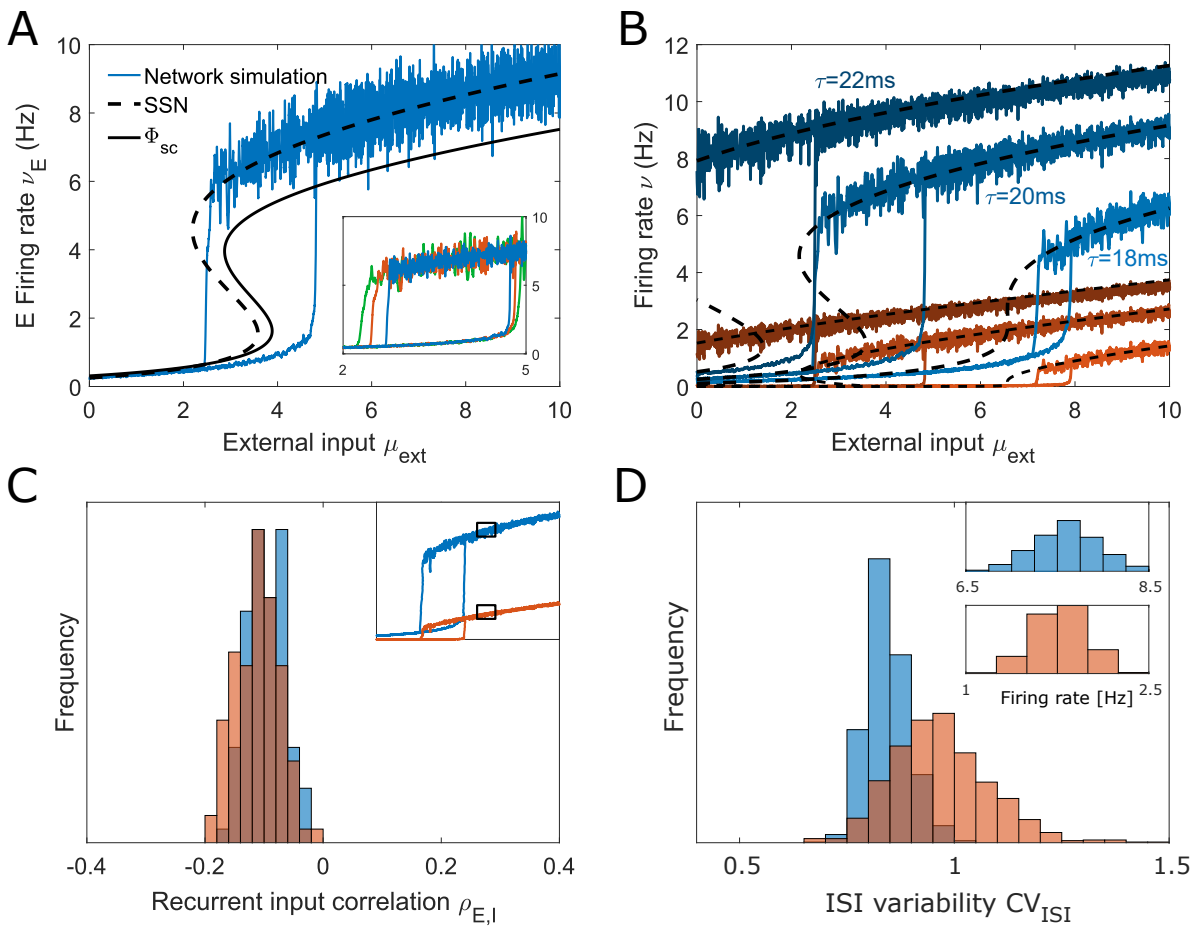


FIG. 4. SSN-predicted bistability and persistent activity can be observed in spiking network simulations. (A) E and I firing rates as a function of external input in a bistable network (coexistence of high and low activity states for a given external input). Simulated LIF spiking activity is in line with the self-consistent solution (Φ_{sc}) and is also accurately predicted by the SSN. The Φ_{sc} -predicted firing rates diverge slightly from the spiking network simulation because of the use of exponential synapses, which lead to correlated recurrent noise (Methods, Appendix). The inset illustrates that the width of the bistability window can vary between simulations of the same network due to the spontaneous transitions between the two states. (B) The width of the bistability window depends on the excitatory membrane time constant: higher values of τ_E lead to broader bistability windows, which are shifted leftward. If τ_E is sufficiently large for the bistability window to exist for zero feedforward input, the network can sustain persistent activity. (C) E/I input correlations to E (blue) and I (red) neurons indicate a persistent but weak E/I correlation, suggesting that the network is only loosely E/I balanced. (D) The coefficient of variation of the interspike intervals (CV_{ISI}) is near 1, which is compatible with a Poisson process and demonstrates that activity is asynchronous and irregular. Both (C) and (D) are measured in the upstate, as shown in the inset to (C). All parameters are given in Table II.

also offer the possibility to store information. Once the network has been switched into a different activity state by a strong perturbation, it remains in the same state even after the perturbation withdrawal.

E. Computational regimes and their position in input space

In previous sections, we demonstrated that the SSN framework can be used to locate specific computational regimes such as supersaturation and the paradoxical effect in parameter space. Here, we focused on the activity

regimes associated with the 2D space of the feedforward input and the I/E external input ratio (μ_{ext}, r). For two examples of connectivity matrices J , we scanned the 2D input space for supersaturation (Eq. B5), ISN (Eq. 7) and bistability using the characteristic function \mathcal{F} as defined in [28] (Eq. B10). We also show the input regimes for which the network permits a balanced limit solution (Eq. B3). Importantly, the sign of the determinant of the weight matrix ($\det J = J_{EI}J_{IE} - J_{EE}J_{II}$) determines the number of SSN steady states (stable and unstable): for $\det J > 0$, the SSN has an odd number of steady states whereas for $\det J < 0$ the number of steady states is even [15, 28]. Thus, networks with positive $\det J$ (Fig. 5A) al-

ways have at least one steady state. In the network shown in Fig. 4, bistability occurs when the system transitions from having one steady state to having three (two stable and one unstable). On the other hand, networks which have a negative $\det J$ (Fig. 5B) can lack steady states at all, and any possible stable steady state coexists with an unstable steady state [28]. Finally, the sign of $\det J$ also determines whether the system can have a stable balanced state (Eq. B3) or lacks it. In networks where the sign of $\det J$ is negative, a balanced state solution can exist with positive firing rates if $r > \max(\frac{J_{II}}{J_{EI}}, \frac{J_{IE}}{J_{EE}})$, but it is unstable [24].

Fig. 5 shows the map of feedforward inputs and the corresponding computational regime for two examples of the connectivity matrix J . Panel A corresponds to the connectivity parameters from the bistable network shown in Fig. 4 with $\det J > 0$. The region where bistability is expected corresponds to the results in Fig. 4A with $r = 1$. The balanced state exists and is stable for low values of r . Panel B corresponds to a network with $\det J < 0$. In this case, the balanced state only exists for high values of r , but it is unstable. Furthermore, we find that for large input and low r , the network does not have a steady state. In this region, inhibition cannot stabilize the network, and the activity blows up. The same analysis is also performed for the supersaturating network in Fig. 2 and the mouse V1 network in Fig. 3A (Fig. S1). Overall, using the SSN model, we can precisely locate the regions corresponding to distinct behaviors of spiking networks in their parameter space. Notably, we observe that the sign of the determinant of the connectivity matrix J plays a crucial role in the type of activity regimes available to the network (Appendix).

F. Effect of network size on network response nonlinearity

While medium-sized networks can generate diverse nonlinear responses to external input, the balanced state framework implies that network response becomes linear as network size approaches infinity. How do networks transition from nonlinear to linear regimes for increasing network size N ? To tackle this question, we re-scaled the recurrent connections J_{XY} by the factor $1/\sqrt{N}$ as a function of network size N , and increased N from $N = 4 \times 10^3$ to 5×10^5 . This parameter re-scaling follows the convention of the balanced state theory [2, 16] and allows us to address whether these nonlinear spiking networks converge to the expected balanced state, and if so, when and how.

A dynamically stable balanced state limit can only exist if $\det J$ is positive and the fraction of external input weights r satisfies $0 < r < \min(\frac{J_{II}}{J_{EI}})$, see Eq. B3. In our network convergence study, we focus on three spiking networks: one supersaturating network with $\det J > 0$ (shown in Fig. 2), one bistable network with $\det J > 0$ (shown in Fig. 4), and a supersaturating network with

$\det J < 0$ (presented in Fig. 5B, where we set $r = 3$). Among our three example networks, we have one example for which a balanced state limit does not exist (supersaturation), one network with a stable balanced state solution (bistable network) and one network with a balanced state that exists but is dynamically unstable ($\det J < 0$). In all three cases, the self-consistency solution Φ_{sc} remains an accurate description of the spiking network mean activity across different network sizes N (Fig. 6A, C, E), and its predictions align qualitatively and quantitatively with the SSN model (Fig. 6B, D, F, inset).

We find that the network response can remain nonlinear even for very large network sizes consisting of up to half a million neurons (see inset). By re-scaling the feedforward input with the factor $1/\sqrt{N}$ [2, 16], the network response should converge toward a single linear solution - the balanced limit (Fig. 6B, D, E). In the case of the supersaturating network, the balanced limit does not exist, as it would lead to negative E firing rates. Therefore, in the limit of large network size the E firing rate tends to zero. In the case of the bistable network, a balanced state limit does exist but the network response is still far from converging to it, even for $N = 5 \times 10^5$. Finally, for the network with $\det J < 0$ and $r = 3$, the network exhibits supersaturation for $N = 4000$ (see Fig. 5B). However, as N increases, the network enters a region for which there is no steady state, and where the firing rates blow up. This behaviour is observed in spiking network simulations, the mean field Φ_{sc} and the SSN solution. Interestingly, the LIF simulation appears to be more stable than the mean-field predictions, and for $N = 20000$ the firing rates grow substantially but do not explode as it is the case for the mean-field solution. The inset shows how this instability is caused by the collision of the two steady state branches, leaving a gap where the firing rates are unbound. For this network, the mean-field solution converges to the balanced limit as N increases (Fig. 6F). However, the balanced state limit is unstable here, and it only matches with the unstable mean-field solution (high firing rates part of the branch) whereas the stable low activity solution tends to zero.

Overall, our example networks illustrate that for many classes of spiking networks with biologically plausible sizes and connectivity configurations, the activity will escape the predictions of the balanced state. Depending on the parameters, the balanced state may not exist or be unstable. Yet, even when a stable balanced state exists, it is not guaranteed that it provides a realistic description of network activity, even for unrealistically large sizes N . Next, we investigated whether the correlations between recurrent I and E inputs that accompany the balanced state can also be observed in our networks that are nonlinear and do not conform to the balanced predictions. We have shown that the correlation between E and I recurrent inputs $\rho_{E,I}$ are weak in the bistable network of 4000 neurons (Fig. 4C), which is an indicator that the network is operating far from balance. However, we found that $\rho_{E,I}$ can be network size dependent: increas-

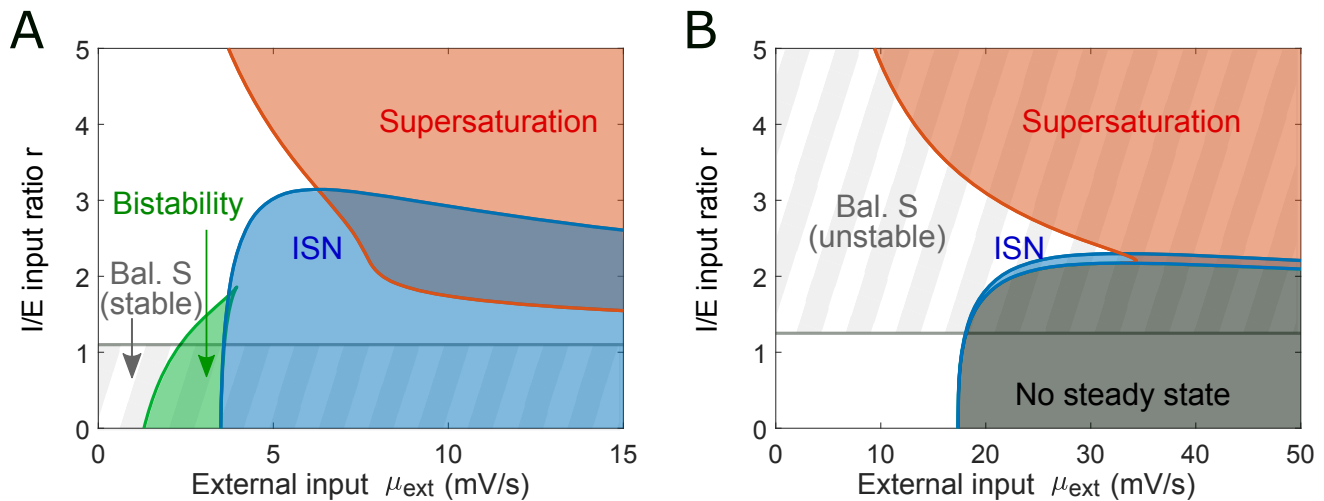


FIG. 5. **Mapping the computational states in the SSN model for two representative connectivity regimes.** (A) Varying the ratio of the external input weights and the amplitude of external drive in a network with a positive $\det J$ allows to traverse different computational regimes (J as in Fig. 4, see parameters in Table II). Gray stripes denote the input space subset with a stable balanced state limit ($N \rightarrow \infty$) which does not exist above the horizontal gray line. The blue area represents the inhibition stabilized regime (ISN). The red area denotes the phase space occupied by supersaturating spiking activity. The green area corresponds to a bistable region (as shown in Fig. 4A with $r=1$). Within the green region, the up-state is in the ISN whereas the down-state is not. We note that the inhibitory stabilization and supersaturation can co-exist. (B) The same analysis is performed on a network with negative $\det J$ (Table II). In this case the balanced state limit is unstable. The ISN region is narrower and there is a broad range of inputs for which the network does not have a steady state solution.

ing the size N from 4000 to 40000 strongly increased the correlation even though the firing rate activity in both cases does not conform to the balanced state solution (Fig. S2). In summary, the observation of strong E-I correlation does not guarantee that the balanced state framework is applicable to predict the firing activity.

III. DISCUSSION

Understanding the activity regimes of biologically sized spiking networks is critical to make sense of experimentally recorded data. The state-of-the-art experimental techniques now enable simultaneous recordings of thousands of neurons [60, 61]. Therefore, it is important to relate experimentally recorded network activity to the computational capabilities of similarly sized networks *in silico*. On the theoretical side, many commonly used models operate in the limit of infinitely large networks (balanced state) [2, 21–24]. Some features predicted by the balanced state, such as high correlations between excitatory and inhibitory currents within neurons, have been observed experimentally [62, 63]. However, other features, such as strong feedforward inputs, have not been experimentally reported, and some experiments argue against it [22, 64–67]. Unlike the balanced state framework, which assumes a tight E/I input balance, alternative models such as the SSN only assume loose E/I balance to achieve inhibitory stabilization [20, 29, 48, 57, 58]. Yet, clarifying under which conditions the predictions of

these two models are conflicting or aligning and how they compare to spiking network simulations has so far remained an open question.

Here, we mapped the computational regimes of spiking networks containing a few thousand neurons across the input space. We chose this size because neuroanatomical and connectomics studies of different cortical regions such as the somatosensory cortex [68–72] indicated that a functional unit such as a minicolumn could contain a few thousand neurons. We also set the range of firing rate activity such that it meets the experimentally reported range of a few Hz [30–37].

We have shown that the nonlinear behavior of medium-sized spiking networks can be quantitatively and qualitatively understood using the SSN model which is based on the power-law activation function of individual neurons. Importantly, the mapping we propose can address the broad activity and connectivity regimes, including those located outside the validity domain of the balanced network formalism. We delineated connectivity regimes where bistability, supersaturation, inhibitory stabilization, or even the absence of steady states can occur in spiking networks. We found that networks can be inhibition stabilized in conditions where a balanced limit does not exist, even though both require strong inhibitory feedback. The ISN can overlap with supersaturation or can be achieved in networks with $\det J < 0$ (Fig. 5), which is outside of the domain of the balanced state framework. Furthermore, we found that network parameters obtained from experimental recordings are incom-

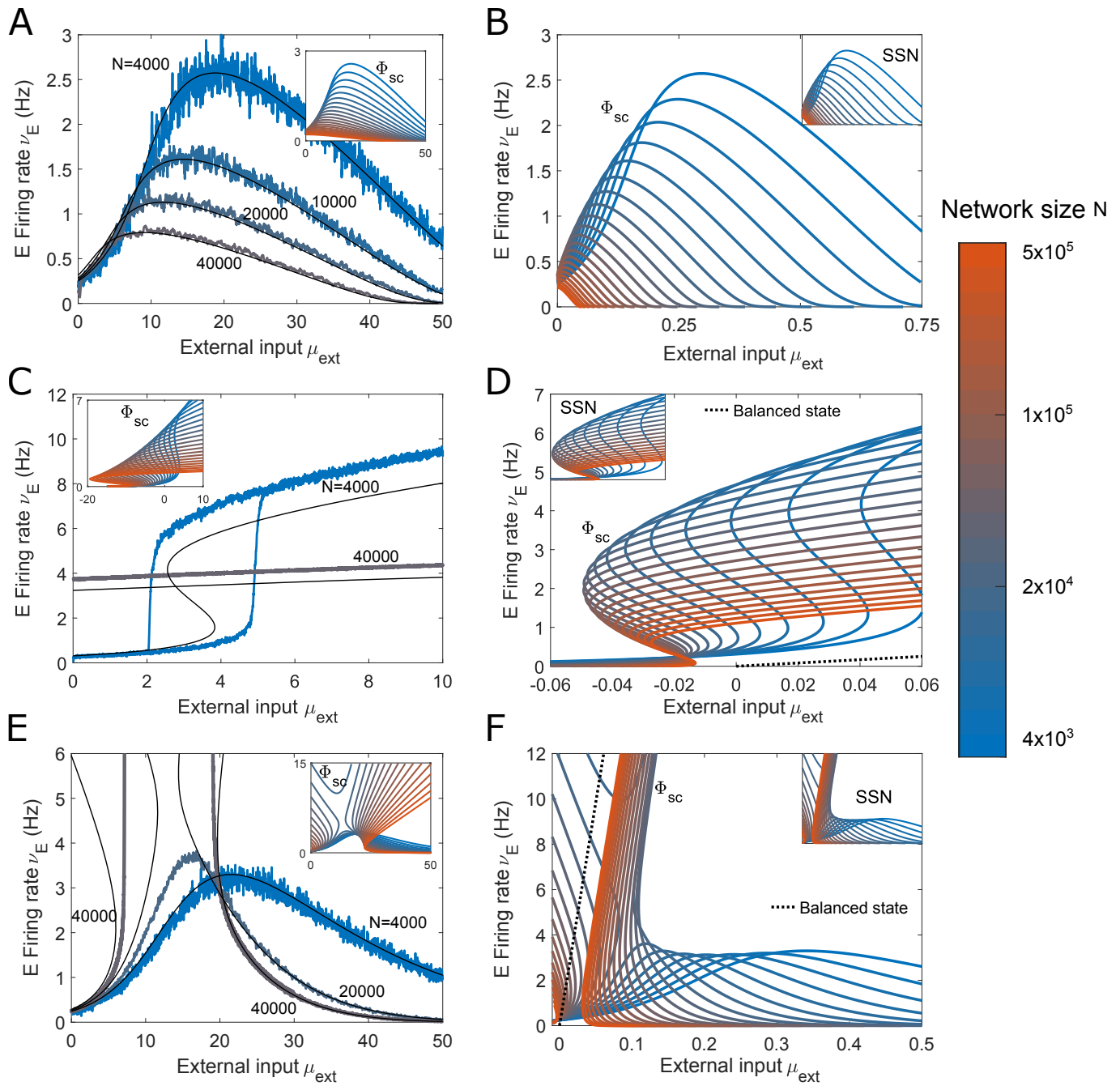


FIG. 6. Increasing network size does not guarantee convergence to a balanced state (A-B) supersaturating network with $\det J > 0$, (C-D) bistable network with $\det J > 0$, (E-F) supersaturating network with $\det J < 0$, parameter regimes which we identified using the SSN framework and studied for $N = 4000$ in previous figures. Here, we gradually increase the size of these networks N and follow the balanced network convention to rescale the weights J_{ij} by \sqrt{N} as the network grows. (A) depicts the excitatory firing in the supersaturating network from Fig. 2A across different network sizes. The colored lines represent spiking networks (from 4×10^3 to 4×10^4), the black lines represent the corresponding mean-field solution Φ_{sc} (Eq. 12). The inset shows the excitatory rate for network sizes from $N = 4 \times 10^3$ (blue) to $N = 5 \times 10^5$ (red), in steps of $\times 10^{0.1}$ obtained using Φ_{sc} . (B) depicts the same network as in A, but the external input is rescaled as $\mu_{\text{ext}}/\sqrt{N}$. This network does not have a balanced state solution (see first condition of Eq. B3). As N grows, the excitatory activity peak becomes smaller and in the limit of very large networks, the excitatory population remains silenced ($\nu_E = 0$). The inset shows that the SSN and Φ_{sc} predict the same behavior as N increases. (C) shows the excitatory activity for the bistable network from Fig. 4, for $N = 4 \times 10^3$ and $N = 4 \times 10^4$. The spiking activity of the spiking LIF network (colored lines) is captured by Φ_{sc} (black lines). As N increases we observe a broadening of the bistability window and a decreasing firing rate, see inset. (D) depicts the same

FIG. 6. network as in panel C but now with rescaled external inputs. The balanced state predicts a linear solution for $N \rightarrow \infty$ limit (dashed line). The convergence to the balanced limit is very slow, and even at $N = 5 \times 10^5$ neurons, the network rates still do not converge to the balanced state. (E) Activity of a spiking network with a negative $\det J$ from Fig. 5B, with $r=3$. As N grows, the excitatory activity dissociates into two distinct branches separated by an unstable region where the firing rates diverge to ∞ . The inset illustrates that this separation occurs when the stable and unstable steady states collide. (F) depicts the same network as in E, now with rescaled external inputs. The balanced solution of this network (dashed line) is unstable (Eq. B3). The unstable solution of Φ_{sc} approaches the balanced state, while the stable solution tends to 0 as N increases.

patible with a balanced solution unless the E population receives a stronger feedforward input than the I population.

We used our theory to target a 2D firing rate trajectory of interest within the spiking neural circuit and could implement an XOR gate by exploiting the intrinsic nonlinearity of the neuronal transfer function (Fig. 2E,G). This suggests that designing a biologically realistic spiking network to perform a task can be facilitated by understanding the input-output relationship in networks.

Studying the network responses across different network sizes, we found that convergence to the balanced state was not always guaranteed, even for unrealistically large network sizes. Activity deviations from the predicted balanced solutions were significant and could be observed even for large E/I correlations of the recurrent inputs. This indicates that strong E/I correlations can be present in networks that do not meet the balanced network predictions on the activity level (e.g., exhibit bistability, Fig. S2). At the same time, networks can be asynchronous irregular in the absence of strong E/I correlations (e.g., Fig. 2). Importantly, we showed that the SSN model could accurately describe spiking networks that do not have a stable balanced limit.

Overall, the balanced state framework has several limitations such as a limited domain of validity, a linear response function, the need for extensive network sizes, and strong feedforward inputs. We show that we can avoid these limitations by using a different rate model, the SSN which assumes a power law transfer function. As we show here, these limitations can be addressed in medium-sized spiking networks by mapping them to the SSN model, which supports nonlinear responses for a broad range of connectivity configurations without the requirement of strong feedforward inputs. It should be noted, however, that other works are addressing these limitations within the balanced state framework, by expanding it. For example, balanced networks with short-term synaptic plasticity have been proposed to permit the emergence of nonlinear activity, such as bistability [21]. Likewise, the experimentally reported small feedforward input which drives spiking activity *in vivo* [64–67] was inconsistent with the original balanced state predictions but was accommodated via the inclusion of broad synaptic weight distributions [22]. Similarly, semi-balanced networks were proposed [23], where neurons which receive net inhibition remain silent. This generated a piecewise-linear manifold which can operate as a nonlinear decision boundary and allowed for a broader

domain of validity than the classical balanced framework. Let us also mention that [27] considered a neuronal transfer function which is nonlinear at onset (noise driven), saturating at high rates (due to a refractory period) and linear in between. This study showed quantitative differences in a set of nonlinear responses of SSN and spiking network models which we resolved using a precise mapping between SSN solutions and spiking network activity.

Our modeling approach considers two homogeneous neural populations, which operate at equilibrium. This approach led to a model with tractable equations and only a few parameters. The complementary modeling strategy, which aims to provide a highly detailed description of a neural system [73] achieves precise biological realism at the expense of mathematical tractability. Future studies could expand the results presented here by including additional features into the network and studying their impact. Such features could include synaptic plasticity which would provide an additional source of network nonlinearity and lead to an even richer repertoire of operational regimes. Similarly, future work could analyze the dynamical properties of spiking networks, as they have been shown to affect the stability of fixed points or to be linked with activity regimes such as oscillations observed in the SSN model [28] (τ_{P_E} and τ_{P_I} in Eq. 2). Moreover, the distribution of synaptic weights has been shown to play a significant role in the overall activity regime of spiking networks [22, 74], suggesting that heterogeneity in the properties of neuronal populations could be an important feature of spiking networks which future works could include in the SSN prediction. Finally, our study of medium-sized spiking networks ($\sim 10^3$ neurons) lays the groundwork for the analysis of much larger networks, such as a whole functional area ($\sim 10^5$ neurons [75]). To that end, the large network would be broken down into a "network of networks" where each system unit represents a single E-I network described by the quantitative mapping developed here.

IV. METHODS

A. Power-law approximation of the input-output transformation in a single neuron

We represent the spiking activity of a neuron using the integrate-and-fire model

$$\frac{dV}{dt} = -V/\tau + I. \quad (8)$$

Where V is the membrane potential, τ is the membrane time constant, and I is the input to the neuron. Upon reaching the firing threshold Θ , $V(t)$ is reset to V_R . If we assume the input I to be white noise with a mean μ and variance σ^2 , the firing rate of the neuron in Eq. 8 can be described by the Ricciardi function Φ [39, 76]

$$\nu = \Phi(\mu, \sigma, \tau) = \left(\tau \sqrt{\pi} \int_{\frac{V_R - \mu \tau}{\sigma \sqrt{\tau}}}^{\frac{\Theta - \mu \tau}{\sigma \sqrt{\tau}}} e^{z^2} (1 + \operatorname{erf}(z)) dz \right)^{-1}. \quad (9)$$

For low inputs, Φ is a supralinear function of the mean input μ and can be accurately approximated by a power law with an exponent $n > 1$ (see Fig. 1). For high inputs, however, Φ becomes linear. In this work, we restrict ourselves to the low firing rate regime ($\nu \leq 10$ Hz) often reported for cortical activity measured *in vivo* [30–37]. In this low activity regime with a constant variance σ and time constant τ , the firing rate can be accurately approximated by a power law Eq. 1, see Fig. 1B. The power law parameters a , b and n are obtained by fitting Eq. 9 (Fig. 1C).

B. LIF spiking network

We consider a spiking network of one excitatory (E) and one inhibitory (I) population with $N_E = \frac{3}{4}N$ and $N_I = \frac{1}{4}N$ LIF neurons, respectively. We assume that both E and I populations are homogeneous, i.e. neurons within each population have the same parameters (membrane time constant τ_X , threshold potential Θ_X , reset value V_{RX}), receive external input with the same mean $\mu_{\text{ext}X}$ and variance $\sigma_{\text{ext}X}^2$, $X \in \{E, I\}$. The E and I populations have different membrane time constants (see black crosses in Fig. 1B), and the feedforward input they receive differs by a factor of r ($\mu_{\text{ext}I} = r\mu_{\text{ext}E}$, see Fig. 1A). Additionally to the feedforward input, the neurons receive recurrent input from other E and I neurons in the network. The connections are randomly generated based on a homogeneous probability of connection, such that each neuron in population X receives inputs from exactly $N_Y p_{XY}$ randomly chosen neurons in population Y , where p_{XY} is the connection probability from population Y to population X . We use two types of synapses, the delta synapse and the exponential synapse.

For delta-synapses, the function

$$I_{XY}(t) = j_{XY} \delta(t - t_s) \quad (10)$$

represents the input from a neuron of the population Y to a neuron in X . Where j_{XY} is the strength of the synapse, t_s is the spike time of the presynaptic neuron, and δ is the Dirac delta function.

In some network configuration, delta synapses promote synchronization of the whole neuronal population. This synchronicity can lead to population spikes [77, 78] which violates the assumption of asynchrony and irregularity in the mean field approach. In order to avoid this synchronization in these cases, we use exponential synapses instead of delta synapses. In exponential synapses, the synaptic potential from a neuron in population Y to a neuron in population X decays exponentially in time

$$I_{XY}(t) = \frac{j_{XY}}{\tau_{sXY}} e^{-\frac{t-(t_s+D)}{\tau_{sXY}}}, \quad t > t_s + D. \quad (11)$$

Where j_{XY} is the strength of the synapse, t_s is the spike time of the presynaptic neuron, τ_{sXY} is the synaptic decay time constant and D is the synaptic delay. This type of synapse prevents synchronization as the effect of each spike is more distributed in time and each synaptic connection has a different delay D .

We use the exponential synapse in the spiking network simulation in Fig. 4, Fig. 6C, Fig. 6E and Fig. S2, and delta synapses in all other cases.

C. Self-consistent network solutions

In this work we derive predictions for the activity regimes of spiking networks using the closed-form solutions offered by the SSN framework (Eq. 3). In some instances, it is useful to compare the SSN predictions to the previously proposed self-consistent network solutions to understand the dynamic origin of the SSN predictions. We provide here the system of self-consistent mean field network equations that arise from the Φ transfer function [78] and that need to be solved numerically to obtain E and I firing rates ν_E and ν_I

$$\begin{cases} \nu_E &= \Phi(\mu_E, \sigma_E) \\ \mu_E &= J_{EE}\nu_E - J_{EI}\nu_I + \mu_{\text{ext}} \\ \sigma_E^2 &= J_{EE}j_{EE}\nu_E + J_{EI}j_{EI}\nu_I + \sigma_{\text{ext}E}^2 \\ \nu_I &= \Phi(\mu_I, \sigma_I) \\ \mu_I &= J_{IE}\nu_E - J_{II}\nu_I + r\mu_{\text{ext}} \\ \sigma_I^2 &= J_{IE}j_{IE}\nu_E + J_{II}j_{II}\nu_I + \sigma_{\text{ext}I}^2 \end{cases} \quad (12)$$

Where J_{XY} is the population-wide connectivity defined by

$$J_{XY} = j_{XY} p_{XY} N_Y. \quad (13)$$

In these equations, the steady state spiking activity of neurons is assumed to follow a Poisson process with a

constant rate ν_X , $X \in \{E, I\}$. The input to each neuron is modeled as a Gaussian process with a mean μ_X , and a variance σ_X^2 . We refer to this approach as "Self-consistency solution" or Φ_{sc} in our figures.

D. Mapping LIF network - SSN

To meet the SSN activity regime with a simulation of spiking LIF neurons, we map the LIF network parameters to SSN parameters. The connectivity parameters J_{XY} in the SSN correspond to the population-wide connectivity defined for the self-consistency solution Φ_{sc} according to Eq. 13. The transfer function parameters a , b and n for each of the populations are obtained by fitting the F-I curve of the neuron obtained with the Φ function Eq. 9, which depends on the LIF membrane time constant τ , reset potential V_R , firing threshold Θ , and the input noise σ (Fig. 1). The noise σ is set to be the external noise σ_{ext} .

In LIF spiking networks, σ_{ext} models the fluctuations in the membrane potential, which can be caused by fluctuations in the external network input as well as originate from intrinsic properties of the neuron [79]. We note that unlike in the Ricciardi mean-field solution Φ_{sc} (Eq. 12), the SSN framework (Eq. 3) does not explicitly model the input noise σ to neurons embedded in a network. Instead, the effect of the noise is implicitly included in the power law approximation of the F-I curve. As a result, the noise in the SSN model is independent of the network activity leading to the assumption that the noise associated with recurrent input is negligible compared to the external noise $\sigma^2 = \sigma_{ext}^2 + \sigma_{rec}^2 \approx \sigma_{ext}^2$. This approximation holds if the firing rates ν and the connection strength j_{XY} are kept low (Eq. 12), which is supported by experimental evidence [6, 30–37] (Appendix).

V. ACKNOWLEDGEMENT

This work was supported by the Max Planck Society, University of Bonn Medical Center, University of Mainz Medical Center, German Research Foundation (CRC 1080 to T.T.) and the Loewe Center for Multiscale Modelling in Life Sciences. We thank Laura Bernandez Timon for comments on earlier versions of the manuscript. We thank Laura Busse, Julijana Gjorgjieva, Jochen Triesch, Simon Renner and all members of our group for scientific discussions. We thank Alexandra Vormberg and Andreas Nold for their contribution at the early stage of the project.

Appendix A: Extraction of experimentally reported network parameters

We use the Cell Types and the Synaptic Physiology databases from the Allen institute [6], to derive biolog-

ically plausible network parameters based on the data collected for layer 2/3 of the mouse visual cortex. Nevertheless, our analyses are not restricted to this brain region, as our framework is applicable to any cortical network. It should be noted that the reported values of network parameters vary largely between sources [8–12]. Furthermore, detailed network simulations using state-of-the-art experimentally measured parameters require optimization of all their recurrent connection weights in order to generate realistic spiking activity [73, 80]. Therefore, we use the parameters for mouse V1 extracted from the Allen institute database [6] as a starting point from which we can explore the range of biologically plausible network connectivity.

We use the Cell Type database [6], to obtain the membrane time constant (τ) as well as the membrane resting and threshold potentials (V_R and Θ), for E and I neurons. The data has been obtained through whole cell patch clamp recording. We use the Synaptic Physiology database [6] to derive the synaptic strength and the probability of connection (j_{XY} and p_{XY}) between E and I neurons. The data has been obtained with octopatching, the simultaneous patch-clamp recording of up to eight neurons. The neurons whose dendrite type is classified as *spiny* constitute the E population. The I population consists of all neurons classified as Vip, PV and Sst.

The reported neuronal properties (τ , V_R and Θ) were obtained by averaging the recorded values from 66 E neurons and 94 I neurons. The probability of connection (p_{XY}) is the fraction of connected pairs with respect to all probed pairs, where 80 EE, 150 EI, 160 IE and 607 II cell pairs were probed, Table I. Finally, we derive the connection strength (j_{XY}) using the peak amplitude (A_{PSP}), the rise time (t_R) and the decay time constant (τ_D) of the postsynaptic potential. These parameters are measured over 66 EE, 46 EI, 3 IE and 29 II connected cell pairs. The mean-field analysis of the network at equilibrium does not depend on the dynamical properties of the synaptic transmission (i.e. the shape of the PSP profile). Instead, the mean-field strength of the synaptic connection is the overall depolarization caused by a single presynaptic spike and it is characterized by a single value j_{XY} , regardless of the dynamics. It is given by the time integral of the synaptic current elicited by one presynaptic spike. We assume that the profile of the postsynaptic membrane potential following a presynaptic spike is a linear increase followed by an exponential decay and that the recorded postsynaptic neuron was at resting potential before receiving its input.

$$PSP(t) = \begin{cases} V_R + A_{PSP} t & \text{if } t < t_R \\ V_R + A_{PSP} e^{-\frac{t-t_R}{\tau_D}} & \text{if } t > t_R \end{cases}$$

Using the LIF equation (Eq. 8) for a single pre-

N_E	3000	
N_I	1000	
p_{EE}	6.5	%
p_{EI}	20	%
p_{IE}	27.5	%
p_{II}	10	%
V_r	0	mV
θ	1	mV
τ_E	20	ms
τ_I	10	ms
σ_{ext}	3	mV/ \sqrt{s}
r	1	

TABLE I. Spiking network parameters used in all figures by default. Parameters deviating from the default ones are specified in figures' panels and captions. For derivation, see section *Extraction of experimentally reported network parameters*.

synaptic spike ($\int Idt = j_{XY}$) yields

$$\begin{aligned} j_{XY} &= \int_0^\infty \left(\frac{dV}{dt} + \frac{V}{\tau} \right) dt \\ &= (V_R - V_\theta) + \frac{\int_0^\infty V dt}{\tau} \\ &= A_{PSP} \frac{t_R/2 + \tau_D}{\tau}. \end{aligned}$$

The synaptic strength is then normalized by the experimentally recorded value of $\Theta - V_R$ [6], so that V_R and Θ can be set to 0 and 1, respectively.

By default, we assume that both populations receive the same external input ($r = 1$). Finally, we assume the input noise σ_{ext} to be 3 mV/ \sqrt{s} , which leads to a power law exponent in the F-I curve close to 3, similar to the values reported in [35, 40] (Table III - E and I populations).

The neuronal parameters used in LIF spiking network simulations are presented in Table I. Fig. 3A is generated with the mouse V1 parameters we extracted from the Allen institute database [6]. We explore the range of biologically plausible network connectivity parameters by modifying the individual connection strengths J_{XY} while keeping them within the range delimited by the lowest and largest experimentally reported connectivity values in mouse V1 ($0.5 \leq J_{XY} \leq 25$). The connectivity parameters used in all network simulations are presented in Table II.

Deriving biologically plausible neural network size for circuit simulations

In our framework, we assume that a network consists of populations of neurons which share a similar external input and preferentially connect together with a homogeneous connection probability and strength. All connections originating from outside this circuit are considered to be feedforward input (see schematic Fig. 1A, bottom).

In biological circuits, it is difficult to determine what can constitute a single network since the brain exhibits a high degree of complexity and does not consist of well-separated circuits. Nonetheless, the analysis of cortical regions in which projection columns can be anatomically identified leads to a consistent order of magnitude for network sizes. Here we set the typical functional network to consist of 3000 E and 1000 I neurons. We present the corresponding citations below.

In the primary visual cortex of mice, it is especially difficult to define a local network size since the cortical map is unstructured, meaning that neurons which share the same receptive field do not co-localize [81]. However, this does not mean that the concept of homogeneous network cannot be applied to mouse V1, as neurons which share the same receptive field preferentially connect together [8]. Since local networks within mouse V1 cannot be defined based on spatial anatomy, we used other regions with anatomically-defined networks to define a reference point.

In mouse somatosensory cortex, distinct neuroanatomical structures known as barrels receive the sensory input from each whisker [4]. These structures are perfect candidates to define the typical scale of a homogeneous network. Their diameter ranges from 100 to 400 μm , with a thickness in layer 4 of 100 μm [4]. Using the neuronal density observed in [5], the number of neurons in these structures ranges from 140 to 2200. In the rat barrel cortex, the number of neurons in each layer of multiple projection columns has specifically been counted [3], and is on the order of $N = 4000$ in L4 and $N = 6000$ in L2/3.

In primates, the primary visual cortex of macaques has been studied extensively. Unlike rodents, the cortical organization of macaque V1 shows a columnar structure, both for eye dominance and orientation preference [82]. Within a range of orientation preference of 10° , such columns are slab-shaped, with a size of 30 μm by 0.5 to 1mm [83]. With the thickness of L2 and L3 being respectively 225 μm and 310 μm [84], and a cell density of 1.3×10^5 neurons/ mm^3 [85]. We can deduce that the number of neurons in L2/3 in such columns is in the range of 1000 to 2000.

In summary, it appears that across species and cortical regions, we can define functional networks with sizes ranging from hundreds to a few thousands of neurons. In this work, we choose a network size of 4000 because it corresponds to the value reported in [3], which is the only study where the neurons in a cortical column were directly counted. In this context, we can use anatomical analyses of neurons in the same brain region to determine the fraction of excitatory and inhibitory neurons in a network [7, 86], which leads to a E/I ratio of 3.5 in layer 2/3. This corresponds to 3110 E neurons and 890 I neurons, which we round to 3000 and 1000 respectively.

Finally, we verify that the network size we assume (i.e. 4000 neurons) is plausible for mouse V1, since we could not use a spatial anatomy feature to define a network. We use the reported probability profile (p) for E \rightarrow E con-

	j_{EE}	j_{EI}	j_{IE}	j_{II}	J_{EE}	J_{EI}	J_{IE}	J_{II}
Mouse V1 Fig. 3A Fig. S1B	3.59	68.1	28.7	108	0.672	13.2	23.7	11.8
Fig. 2 Fig. S1A Fig. 6A-B	10.3	60	7.27	10	2	12	6	1
Fig. 4 Fig. 5A Fig. S2 Fig. 6C-D	25.6	50	8.49	110	5	10	7	11
Fig. 5B Fig. 6E-F	6.41	-5	1.21	-12.5	1.25	1	1	1.25
Fig. 3B	25.4	68.1	28.7	108	4.75	13.2	23.7	11.8

TABLE II. Network connectivity parameters used in all figures. The synaptic weights j_{XY} correspond to the strength of a single spike in spiking networks and are given in μV . The connection strength at the population level J_{XY} , used in mean field solutions, are given in mV (Eq. 13). For Fig. S2 and Fig. 6, synaptic weights are N-dependant since the $1/\sqrt{N}$ scaling is applied; the given weights correspond to $N = 4000$. See section *Extraction of experimentally reported network parameters*.

SSN parameters	a	b	n
Excitatory population	1.08×10^{-4}	-11.1	3.08
Inhibitory population	2.21×10^{-6}	4.8	3.82
Fig. 1 ($\sigma = 1 \text{ mV}/\sqrt{s}, \tau = 20 \text{ ms}$)	3.91×10^{-2}	31.6	2.15
Fig. 1 ($\sigma = 3 \text{ mV}/\sqrt{s}, \tau = 15 \text{ ms}$)	3.23×10^{-5}	-6.4	3.31
Fig. 1 ($\sigma = 5 \text{ mV}/\sqrt{s}, \tau = 10 \text{ ms}$)	4.90×10^{-8}	-63.0	4.20
Fig. 4B ($\sigma = 3 \text{ mV}/\sqrt{s}, \tau = 18 \text{ ms}$)	7.89×10^{-5}	-9.5	3.14
Fig. 4B ($\sigma = 3 \text{ mV}/\sqrt{s}, \tau = 22 \text{ ms}$)	1.60×10^{-4}	-12.1	3.01

TABLE III. SSN parameters used in all figures. The parameters a , b , and n are based on the power-law fit of the LIF F-I curve (Eq. 1), with the input μ in mV/s and the firing rate ν in Hz.

nections as a function of distance for mouse V1 L2/3, Fig. 4B in [54] (near 20% for nearby pairs, going down to 0% for neuron pairs $150\mu\text{m}$ apart). Following this observation, the number of excitatory synapses to an excitatory neuron can be obtained through a 3D spatial integration of this connection probability profile: $n_c = \int \int \int p(\sqrt{x^2 + y^2 + z^2}) \eta dx dy dz$, where η is the density of excitatory neurons. Over an infinite 3D space, we obtain the total number of E→E connections: $n_c = 8.7 \times 10^5 \eta$. Using a neuronal density of 1.64×10^{-4} neurons/ μm^3 [87], and assuming a E/I ratio of 3.5, we obtain $n_c \approx 111$ connections. The result of this rough calculation is in the same order of magnitude as the 195 connections we obtain with 4000 neurons and a probability of connection of 6.5% (as described in [6]), which suggests that this network size is a valid approximation for mouse V1 as well. It should be noted that the network we define does not constitute a single block of cortex due to the salt-and-pepper organization of this brain region, but consists instead of distant neurons which receive the same external input and are homogeneously connected within the network.

Appendix B: Derivation of parameter conditions for specific computational regimes

1. Balanced state framework

Here, we provide for completeness the solutions of the balanced state framework derived previously [16] which we used as a reference to study the convergence to the balanced state. In this framework, the mean input to each population vanishes as the number of neurons in the network increases, $\mu_E \approx \mu_I \approx 0$ and the equations for μ_E and μ_I in Eq. 12 can be simplified to

$$\begin{aligned}\mu_{\text{ext}} &= J_{EI}\nu_I - J_{EE}\nu_E \\ r \times \mu_{\text{ext}} &= J_{II}\nu_I - J_{IE}\nu_E.\end{aligned}$$

The solution of the balanced state equations reads

$$\begin{aligned}\nu_E &= \frac{rJ_{EI} - J_{II}}{J_{EE}J_{II} - J_{EI}J_{IE}} \mu_{\text{ext}} \\ \nu_I &= \frac{rJ_{EE} - J_{IE}}{J_{EE}J_{II} - J_{EI}J_{IE}} \mu_{\text{ext}}.\end{aligned}$$

The balanced state solution is only valid if both ν_E and ν_I are positive for positive input, which corresponds to

the following condition on connectivity

$$\begin{cases} \frac{rJ_{EI} - J_{II}}{J_{EE}J_{II} - J_{EI}J_{IE}} > 0 \\ \frac{rJ_{EE} - J_{IE}}{J_{EE}J_{II} - J_{EI}J_{IE}} > 0. \end{cases} \quad (\text{B1})$$

Stability of the balanced state

Within the balanced state framework, we assume that the change in firing rate of each population is a function of the excess input it receives.

$$\begin{aligned} \frac{d\nu_E}{dt} &= f(\mu_E) \\ \frac{d\nu_I}{dt} &= g(\mu_I). \end{aligned}$$

The firing rate of a population is at the steady state when its total input is balanced ($f(0) = 0$ or $g(0) = 0$). By linearising around a steady state, we get

$$\begin{aligned} \nu_E &= \nu_E^0 + \delta\nu_E \\ \nu_I &= \nu_I^0 + \delta\nu_I \end{aligned}$$

$$\begin{aligned} \nu_E^0 + \delta\dot{\nu}_E &= f(\mu_E^0 + \delta\mu_E) \\ \nu_I^0 + \delta\dot{\nu}_I &= g(\mu_I^0 + \delta\mu_I) \end{aligned}$$

$$\begin{aligned} \delta\dot{\nu}_E &= f'(\mu_E^0)\delta\mu_E = f'(\mu_E^0)(J_{EE}\delta\nu_E - J_{EI}\delta\nu_I) \\ \delta\dot{\nu}_I &= g'(\mu_I^0)\delta\mu_I = g'(\mu_I^0)(J_{IE}\delta\nu_E - J_{II}\delta\nu_I) \end{aligned}$$

This can be rewritten as

$$\begin{bmatrix} \delta\dot{\nu}_E \\ \delta\dot{\nu}_I \end{bmatrix} = \begin{bmatrix} f' & 0 \\ 0 & g' \end{bmatrix} \begin{bmatrix} J_{EE} & -J_{EI} \\ J_{IE} & -J_{II} \end{bmatrix} \begin{bmatrix} \delta\nu_E \\ \delta\nu_I \end{bmatrix}.$$

Where f' and g' are positive (excess input drives the firing rate up), the state (ν_E^0, ν_I^0) is stable if the two eigenvalues of the Jacobian matrix

$$\begin{bmatrix} f'J_{EE} & -f'J_{EI} \\ g'J_{IE} & -g'J_{II} \end{bmatrix}$$

have negative real parts.

The eigenvalues λ_1 and λ_2 are roots of the polynomial

$$\lambda^2 - \lambda(f'J_{EE} - g'J_{II}) + f'g'(J_{EI}J_{IE} - J_{EE}J_{II})$$

or

$$\begin{aligned} \lambda_1, \lambda_2 &= \left(\frac{f'J_{EE} - g'J_{II}}{2} \right) \\ &\times \left(1 \pm \sqrt{1 - \frac{4f'g' \det J}{(f'J_{EE} - g'J_{II})^2}} \right). \end{aligned}$$

The steady state is stable \iff

$$\begin{cases} f'J_{EE} - g'J_{II} < 0 \\ \det J > 0 \end{cases} \quad (\text{B2})$$

The first condition requires that the response of the inhibitory population (g') is fast and strong enough to prevent a runaway excitatory feedback loop. We do not use this condition here because it depends on the dynamic properties of the network (f and g functions), which are beyond the scope of this work. However, the second condition constrains the connectivity matrix such that $J_{EI}J_{IE} - J_{EE}J_{II} > 0$ [24].

The condition on the existence of a non-negative balanced state (Eq. B1) can be combined with the stability condition on connectivity (Eq. B2) to delineate the parameter range where a balanced state limit exists and is stable [16, 24]:

$$\begin{cases} 0 < r < \min\left(\frac{J_{II}}{J_{EI}}, \frac{J_{IE}}{J_{EE}}\right) \\ \det J > 0 \end{cases} \quad (\text{B3})$$

We use Eq. B3 to visualize the parameter range where a balanced state exists in Fig. 5 and Fig. S1.

2. Condition on supersaturation

Supersaturation is characterized by a decrease in the excitatory firing rate with increasing external input: $\frac{d\nu_E}{d\mu_{\text{ext}}} < 0$.

Linearizing the system around a fixed point, leads to the relation between input and firing rate at the steady state [47]:

$$\begin{bmatrix} \delta\nu_E \\ \delta\nu_I \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \underbrace{\begin{bmatrix} f'_E & 0 \\ 0 & f'_I \end{bmatrix}}_{=F} \underbrace{\begin{bmatrix} J_{EE} & -J_{EI} \\ J_{IE} & -J_{II} \end{bmatrix}}_{=J} \right)^{-1} \times \left(\begin{bmatrix} f'_E & 0 \\ 0 & f'_I \end{bmatrix} \begin{bmatrix} \delta\mu_{\text{extE}} \\ \delta\mu_{\text{extI}} \end{bmatrix} \right).$$

Where $\delta\nu$ and $\delta\mu_{\text{ext}}$ are the firing rates and external inputs linearized around a steady state. The functions f_E and f_I are the input-firing rate transfer functions of the two populations. The functions f'_E and f'_I are the derivatives with respect to input, calculated at the fixed point. The effect of a change of external input yields:

$$\begin{bmatrix} \delta\nu_E \\ \delta\nu_I \end{bmatrix} = \frac{\begin{bmatrix} f'_E(1 + f'_I J_{II}) & -f'_E f'_I J_{EI} \\ f'_E f'_I J_{IE} & f'_I(1 - f'_E J_{EE}) \end{bmatrix}}{|I - FJ|} \begin{bmatrix} \delta\mu_{\text{extE}} \\ \delta\mu_{\text{extI}} \end{bmatrix}. \quad (\text{B4})$$

In particular, the effect of external input on the excitatory firing rate yields

$$\frac{d\nu_E}{d\mu_{\text{ext}}} = \frac{f'_E(1 + f'_I(J_{II} - rJ_{EI}))}{|I - FJ|}$$

As shown in [47], $|I - FJ|$ must be positive for the fixed point to be stable. Furthermore, f'_I and f'_E are assumed to be positive, meaning that the F-I curves are monotonically increasing. This leads to the following condition for supersaturation:

$$f'_I(J_{II} - rJ_{EI}) < -1.$$

Which is only possible if $r > \frac{J_{II}}{J_{EI}}$. In the SSN, the transfer function is a power law (Eq. 1). Its derivative is then $f'_X = n_X a_X^{\frac{1}{n_X}} \nu_X^{\frac{n_X-1}{n_X}}$. This leads to the condition that supersaturation occurs when the fixed point inhibitory firing rate is sufficiently high [15]:

$$\nu_I > n_I^{-1} \sqrt{\frac{1}{a_I(n_I(rJ_{EI} - J_{II}))^{n_I}}} \quad (\text{B5})$$

For completeness, it is worth mentioning that there exists one particular edge-case where $r > \frac{J_{II}}{J_{EI}}$ cannot lead to $\frac{d\nu_E}{d\mu_{\text{ext}}} < 0$. This occurs when the only stable state of the system is such that $\nu_E = 0$ regardless of μ_{ext} . This is discussed in the section *Effect of J_{EE} on the occurrence of ISN* and requires $rb_E - b_I > 0$. Since these cases correspond to situations where recurrent inhibition is strong enough to prevent any excitatory activity, we consider them to be supersaturating.

Modulation of E firing rate peak in supersaturating activity regime

Here, we explain how we modified the height of the E firing rate peak in Fig. 2B. We begin by analyzing how the value of maximal E firing rate depends on SSN parameters for supersaturating networks. A characterizing property of the maximal E firing rate ν_E is that it satisfies $\frac{d\nu_E}{d\mu_{\text{ext}}} = 0$. As shown previously [15], this occurs for $\nu_I = (a_I^{\frac{1}{n_I}} n_I (J_{EI}r - J_{II}))^{-\frac{n_I}{n_I-1}}$, where $J_{EI}r - J_{II}$ must be positive. Since ν_I and ν_E are both positive, we can remove the $(\cdot)_+$ operator in the SSN equations and apply the inverse power law exponents to both sides to express μ_{ext} Eq. 3

$$\begin{aligned} a_E^{-\frac{1}{n_E}} \nu_E^{\frac{1}{n_E}} - (J_{EE}\nu_E - J_{EI}\nu_I - b_E) &= \mu_{\text{ext}} \\ r^{-1} a_I^{-\frac{1}{n_I}} \nu_I^{\frac{1}{n_I}} - r^{-1} (J_{IE}\nu_E - J_{II}\nu_I - b_I) &= \mu_{\text{ext}}. \end{aligned}$$

We combine two above equations to obtain

$$\begin{aligned} a_E^{-\frac{1}{n_E}} \nu_E^{\frac{1}{n_E}} + \left(\frac{J_{IE}}{r} - J_{EE}\right)\nu_E &= \\ - (J_{EI} - r^{-1}J_{II})\nu_I + r^{-1} a_I^{-\frac{1}{n_I}} \nu_I^{\frac{1}{n_I}} - b_E + \frac{b_I}{r}. \end{aligned} \quad (\text{B6})$$

Substituting $\nu_I = (a_I^{\frac{1}{n_I}} n_I (J_{EI}r - J_{II}))^{-\frac{n_I}{n_I-1}}$ in the above equation, we obtain

$$\begin{aligned} a_E^{-\frac{1}{n_E}} \nu_E^{\frac{1}{n_E}} + \left(\frac{J_{IE}}{r} - J_{EE}\right)\nu_E &= \\ \frac{1}{r} (a_I n_I (rJ_{EI} - J_{II}))^{-\frac{1}{n_I-1}} \left(1 - \frac{1}{n_I}\right) - b_E + \frac{b_I}{r}. \end{aligned} \quad (\text{B7})$$

The solution of Eq. B7 corresponds to the maximal E firing rate in the supersaturating activity regime. To increase the E firing rate peak, we modified r , J_{IE} , and J_{EI} . Specifically, we decreased r and modified J_{IE} and J_{EI} such that the terms J_{IE}/r and $J_{EI}r$ remained constant. If $J_{IE}/r - J_{EE}$ is positive, the left-hand side of the equation is a monotonically increasing function of ν_E . As $1/r$ increases, the right side of Eq. B7 moves upward and the corresponding ν_E on the left side must increase as well. In this way the unique solution ν_E of Eq. B7 - the maximal E firing rate - increases as $1/r$ increases. The method demonstrated here assumes that the left hand side of Eq. B7 is an increasing function of ν_E . This is the case if $J_{IE}/r - J_{EE}$ is positive (as in Fig. 2), and the peak of the excitatory activity can be increased infinitely. On the other hand, if $J_{IE}/r - J_{EE}$ is negative, the peak of supersaturation is bounded. In particular, for supersaturating networks ($J_{EI}r - J_{II}$) for which $\det J$ is negative, $J_{IE}/r - J_{EE}$ is negative. The peak of supersaturation is therefore bounded, and decreasing r can lead to an unstable situation where no steady state exist (as illustrated in Fig. 5B).

3. Condition on the paradoxical effect

The paradoxical effect [25, 47, 48] is characterized by a decrease of the I firing rate, as the external input to the I population is increased: $\frac{d\nu_I}{d\mu_{\text{extI}}} < 0$. Here again, the effect of a change in external input is given by Eq. B4 where $|I - FJ|$ and f'_E must be positive. In this case, $\frac{d\nu_I}{d\mu_{\text{extI}}}$ can only be negative if

$$f'_E J_{EE} > 1.$$

This condition is equivalent to the condition for the instability of the excitatory subnetwork [47]. The paradoxical effect is therefore a feature of the inhibition stabilized network (ISN) as it occurs when the activity of the E population is only made stable thanks to the suppression from the I population. Using the SSN transfer function (Eq. 1), the condition on the paradoxical effect leads to Eq. 7

Effect of J_{EE} on the occurrence of ISN

Since the network is inhibition-stabilized for fixed points which satisfy Eq. 7, it seems that any network

can be in the ISN state, granted that J_{EE} is sufficiently strong. Here we ask whether there is any counterexample to this. Are there networks which cannot enter ISN, no matter how high J_{EE} is?

From Eq. B6, we impose that $\nu_E = (a_E n_E^{n_E} J_{EE}^{n_E})^{-\frac{1}{n_E-1}}$, which corresponds to the onset of inhibition-stabilization to get

$$\left(\frac{\nu_I}{a_I}\right)^{\frac{1}{n_I}} - (rJ_{EI} - J_{II})\nu_I = \frac{1}{(a_E n_E J_{EE})^{\frac{1}{n_E-1}}} \times \left(r - \frac{r}{n_E} + \frac{J_{IE}}{n_E J_{EE}}\right) + rb_E - b_I. \quad (\text{B8})$$

This equation can be seen as the crossing of two functions where the left hand side f is a function of the I activity ν_I and the right-hand side g is independent of ν_I but varies with J_{EE} .

$$f(\nu_I) = g(J_{EE}).$$

If there is a value of ν_I which satisfies this equation, the network has a fixed point such that ν_E is at the onset of ISN. We note from Eq. B6 that

$$\frac{df}{d\nu_E} > 0 \iff \nu_E < a_E \left(a_E n_E \left(J_{EE} - \frac{J_{IE}}{r} \right) \right)^{-\frac{n_E}{n_E-1}}.$$

Since the value of ν_E at which we operate in Eq. B8 is smaller than this (for $J_{IE} \neq 0$), the network cannot be in the ISN if f and g do not intersect and f remains below g .

The f function starts at 0 for $\nu_I = 0$ and is either monotonously increasing (if $J_{II} > rJ_{EI}$) and tends to ∞ , or it reaches a maximum and then tends to $-\infty$ (if $rJ_{EI} > J_{II}$). The second case corresponds to supersaturation, where high values of ν_I can be sustained with low recurrent excitation. Whatever the parameters, f first rises and always has at least some positive values.

The g function can be represented as a polynomial

$$g(X) = \frac{J_{IE}}{n_E (a_E n_E)^{\frac{1}{n_E-1}}} X^{n_E} + \frac{r(n_E - 1)}{(a_E n_E)^{\frac{1}{n_E-1}}} X + rb_E - b_I.$$

Where $X = J_{EE}^{-\frac{1}{n_E-1}}$. Since the coefficients in front of X^{n_E} and X are both positive and X is necessarily positive (because J_{EE} is positive), g can take all values larger than $rb_E - b_I$. This means that, with J_{EE} as a free parameter, equation Eq. B8 will have a solution if there is any ν_I such that $f(\nu_I) > rb_E - b_I$.

For values of b_E and b_I such that $rb_E - b_I < 0$, the network can always enter an inhibition-stabilized state by tuning J_{EE} . Interestingly, this is always the case with the values of b_E and b_I we obtained from fitting the F-I

curve (See Table III) because experimentally reported neuronal parameters [6] are such that $\tau_E > \tau_I$. Similarly, in networks which do not satisfy the supersaturation condition ($J_{II} > rJ_{EI}$), g always crosses f .

On the other hand, in networks for which f has a maximum ($J_{II} > rJ_{IE}$), if $rb_E - b_I$ is higher than this maximum, the functions f and g never cross regardless of the value of J_{EE} . In this scenario, the network can never reach a steady state where it is inhibition-stabilized. These cases correspond to situations where the E activity is always suppressed and there is no stable steady state with $\nu_E > 0$. If J_{EE} is large, another unstable steady state will exist (for $\det J < 0$, see section *Parity of solutions* below), and if the system is perturbed enough to reach it the E activity will enter an unlimited feedback loop which leads to always increasing activity. These systems cannot be in the ISN regardless of J_{EE} because the E activity is either entirely silent or cannot be stabilized by inhibition.

4. Condition on multiplicity of solutions

As shown in [28], the two-dimensional SSN equation (Eq. 3) can be rewritten as a single characteristic function \mathcal{F} , where the steady states of the system correspond to zeros of \mathcal{F} .

$$\mathcal{F}(\mu_E) = a_E J_{EE} (\mu_E - b_E)_+^{n_E} - a_I J_{EI} (\mu_I(\mu_E) - b_I)_+^{n_I} - \mu_E + \mu_{\text{ext}}. \quad (\text{B9})$$

Where μ_I is a function of μ_E : $\mu_I(\mu_E) = \frac{a_E \det J}{J_{EI}} (\mu_E - b_E)_+^{n_E} + \frac{J_{II}}{J_{EI}} \mu_E + \mu_{\text{ext}} \left(r - \frac{J_{II}}{J_{EI}} \right)$. The first term of \mathcal{F} is zero if the E population is silenced ($\nu_E = 0$), and the second one if the I population is silenced ($\nu_I = 0$). For any value μ_E which satisfies $\mathcal{F}(\mu_E) = 0$, the corresponding excitatory firing rate is given by the power law F-I function Eq. 1.

The number of zero crossings of the function \mathcal{F} corresponds to the number of fixed points of the system. Since \mathcal{F} is a continuous function, its number of zero crossings only changes when two solutions merge into one or when one solution splits into two. This corresponds to the situation when extrema of \mathcal{F} fall on zero.

$$\begin{cases} \mathcal{F}(\mu_E) = 0 \\ \mathcal{F}'(\mu_E) = 0 \end{cases} \quad (\text{B10})$$

Where \mathcal{F}' denotes the derivative of \mathcal{F} with respect to μ_E . This condition corresponds to changes in the number of solutions. The parameters comprised between these boundaries have the same number of solutions. This approach can be used to delimit the range of bistability or absence of solutions (as shown in Fig. 5). Within such a region, the number of network states is obtained by determining the number of zero crossings of \mathcal{F} .

Parity of solutions

The number of zero crossings of the \mathcal{F} function can be studied through its limits. Assuming that the F-I curves are supralinear ($n_E > 1$ and $n_I > 1$), we get:

$$\begin{aligned} \lim_{\mu \rightarrow -\infty} \mathcal{F}(\mu) &= \infty \\ \lim_{\mu \rightarrow +\infty} \mathcal{F}(\mu) &= \lim_{\mu \rightarrow +\infty} a_E J_{EE} \mu_E^{n_E} \\ &\quad - a_I J_{EI} \left(\frac{a_E \det J}{J_{EI}} \mu_E^{n_E} \right)_+^{n_I}. \end{aligned}$$

If $\det J > 0$, the second limit tends to $-\infty$. Therefore, the function has at least one zero and the \mathcal{F} function for positive determinants has an odd number of solutions (Mean-value theorem). On the other hand, if the determinant is negative, the second limit tends to $-\infty$. In that case, there is no guarantee that the system has a fixed solution and the number of solutions is even. Multiple roots (Eq. B10) are counted separately in this calculation.

Appendix C: Noise contribution of exponential synapses

In this section we expand the mean-field approximation of the recurrent noise to the case of exponential synapses. In the case of the delta synapse (Eq. 10), by assuming a Poisson spike train, the standard deviation is given by $\sigma_{XY} = j_{XY} J_{XY} \nu_{XY}$ (Eq. 12). For exponential synapses however, since the current is distributed in time Eq. 11, the noise is lower than this. The exponential synapse corresponds to a shot noise process, and its covariance is given by [88, 89]:

$$Cov(i_t, i_{t+\Delta t}) = \frac{jJ\nu}{2\tau_S} e^{-\frac{\Delta t}{\tau_S}}.$$

Over a period Δt , the variance of the received current is

$$\begin{aligned} Var\left(\int_{\Delta t} idt\right) &= \int_0^{\Delta t} \int_0^{\Delta t} Cov(i_t, i_{t'}) dt dt' \\ &= 2 \int_0^{\Delta t} \int_0^{t'} \frac{jJ\nu}{2\tau_S} e^{-\frac{t'-t}{\tau_S}} dt dt' \\ &= jJ\nu\tau \left(\frac{\Delta t}{\tau_S} - \left(1 - e^{-\frac{\Delta t}{\tau_S}}\right) \right). \end{aligned}$$

The Φ equation is derived from the analysis of a Ornstein Uhlenbeck process with white noise [39], so that the noise σ considered in Φ is the rate of increase of the variance. For exponential synapses, we then approximate a white noise Ornstein Uhlenbeck process by using the rate of increase of variance:

$$\sigma_{OU}^2 = \frac{Var\left(\int_{\Delta t} idt\right)}{\Delta t} = jJ\nu \left(1 - \frac{\tau_S}{\Delta t} \left(1 - e^{-\frac{\Delta t}{\tau_S}}\right)\right).$$

Since the current at different time steps is correlated, the variance of the overall cumulative current over a time period Δt is not a linear function of Δt . This contrasts with white noise processes where this rate is a constant.

For $\Delta t \ll \tau_S$, σ_{OU} tends to zero whereas for $\Delta t \gg \tau_S$, σ_{OU} approaches the Poisson limit $jJ\nu$. In the context of the mean field analysis of the network, we set the period Δt over which σ_{OU} is considered to be the period between two successive postsynaptic spikes. Therefore, we use the expected value of σ_{OU} over the distribution of ISI. For Poisson processes, the ISI follows an exponential distribution.

$$\begin{aligned} E[\sigma_{OU}^2] &= \int_0^\infty \sigma_{OU}^2(ISI) P(ISI) dISI \\ &= \int_0^\infty jJ\nu_{pre} \left(1 - \frac{\tau_S}{t} \left(1 - e^{-\frac{t}{\tau_S}}\right)\right) \nu_{post} e^{-\nu_{post}t} dt \\ &= jJ\nu_{pre} \left(1 - \nu_{post}\tau_S \ln\left(1 + \frac{1}{\nu_{post}\tau_S}\right)\right). \end{aligned}$$

All in all, the noise originating from the recurrent exponential synapses can be quantified and the Ricciardi self-consistency solution Eq. 12 can be adapted to take this effect into account:

$$\begin{aligned} \sigma_X^2 &= j_{XE} J_{XE} \nu_E \left(1 - \nu_X \tau_{S_{XE}} \ln\left(1 + \frac{1}{\nu_X \tau_{S_{XE}}}\right)\right) \\ &\quad + j_{XI} J_{XI} \nu_I \left(1 - \nu_X \tau_{S_{XI}} \ln\left(1 + \frac{1}{\nu_X \tau_{S_{XI}}}\right)\right) + \sigma_{extX}^2. \end{aligned}$$

However, it should be noted that this approach can only provide an approximation of the firing of LIF neurons since the Φ function is only exact for uncorrelated white noise input.

Appendix D: Supplemental figures

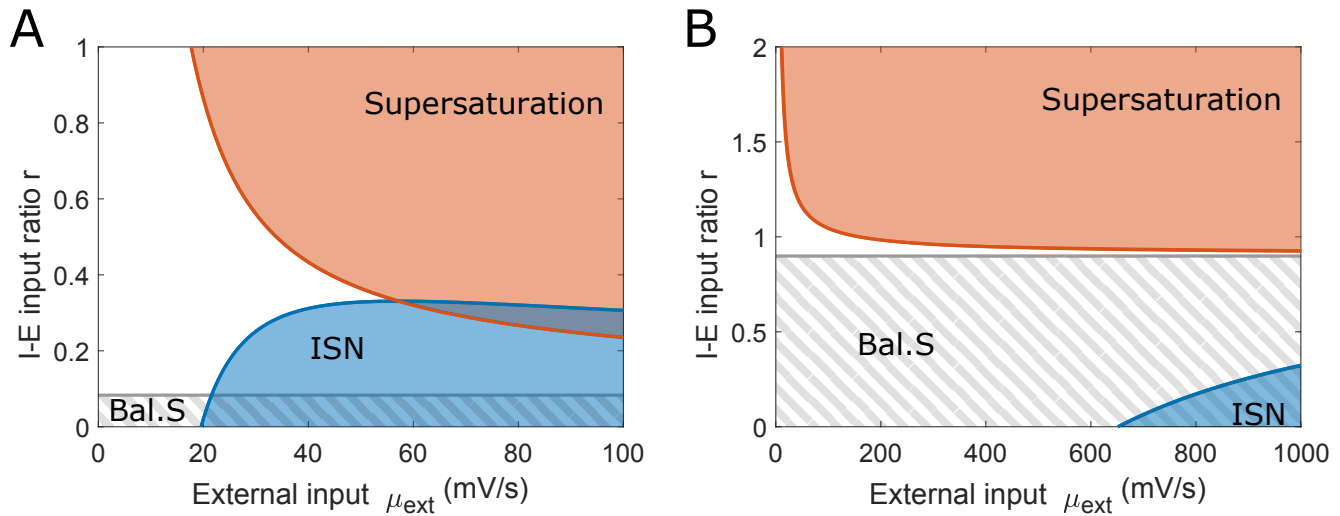


FIG. S1. **Additional map of computational regimes.** These maps are equivalent to the maps shown in Fig. 5, and are generated for the connectivity of the supersaturating network shown in Fig. 2 and the mouse V1 network shown in Fig. 3A (A) The map shows many similarities to the map shown in Fig. 5A. The balanced state is only defined for low r values across external input values. The network can be inhibition stabilized for large input and low r , whereas supersaturation occurs for large input and high r . The supersaturation and ISN regions overlap. However, unlike Fig. 5A, this network does not have a bistable regime in the range of inputs presented here.(B) Compared with the phase space in panel A, the ISN state (blue area) appears more difficult to achieve for this network as it requires much higher external input to reach. We show in Fig. 3B that increasing J_{EE} makes the ISN accessible for external inputs μ_{ext} lower than 100 mV/s.

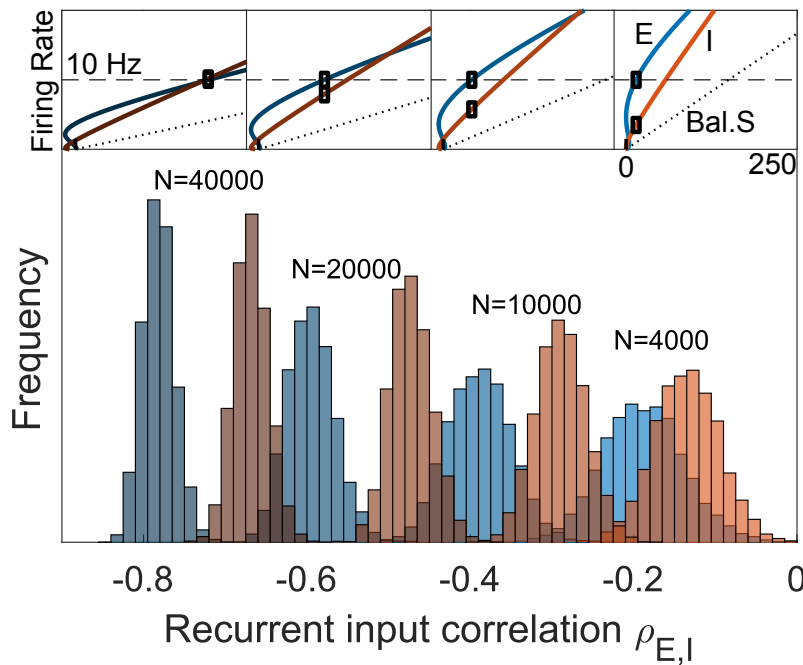


FIG. S2. **Effect of network size on E/I input balance.** As the size of networks increases, the correlation between incoming E and I currents ($\rho_{E,I}$) becomes stronger. This is measured in E (blue) and I (red) neurons for the same bistability case as Fig. 4 and Fig. 6C, at the point at which the excitatory firing rate reaches 10Hz (see inset above, for each network size). This shows that the E-I balance gets tighter as the network size increases even though the firing rates are far from the balanced state limit (dashed line, inset) and the network remains non-linear.

- [1] Daniel J Powell, Eve Marder, and Michael P Nusbaum. Perturbation-specific responses by two neural circuits generating similar activity patterns. *Current Biology*, 31(21):4831–4838, 2021.
- [2] C van Vreeswijk and H Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726, 1996.
- [3] Hanno S Meyer, Verena C Wimmer, Marcel Oberlaender, Christiaan PJ De Kock, Bert Sakmann, and Moritz Helmstaedter. Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissal cortex. *Cerebral cortex*, 20(10):2277–2286, 2010.
- [4] Thomas A Woolsey and Hendrik Van der Loos. The structural organization of layer iv in the somatosensory region (si) of mouse cerebral cortex: the description of a cortical field composed of discrete cytoarchitectonic units. *Brain research*, 17(2):205–242, 1970.
- [5] Alessandro Motta, Manuel Berning, Kevin M Boergens, Benedikt Staffler, Marcel Beining, Sahil Loomba, Philipp Hennig, Heiko Wissler, and Moritz Helmstaedter. Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science*, 366(6469), 2019.
- [6] Allen Institute for Brain Science. Synaptic physiology coarse matrix dataset, 2019. Available from: <https://brain-map.org/explore/connectivity/synaptic-physiology>.
- [7] Henry Markram, Eilif Muller, Srikanth Ramaswamy, Michael W. Reimann, Marwan Abdellah, Carlos Aguado Sanchez, Anastasia Ailamaki, Lidia Alonso-Nanclares, Nicolas Antille, Selim Arsever, Guy Antoine Atenekeg Kahou, Thomas K. Berger, Ahmet Bilgili, Nenad Buncic, Athanassia Chalimourda, et al. Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, 163(2):456–492, October 2015.
- [8] Ho Ko, Sonja B. Hofer, Bruno Pichler, Katherine A. Buchanan, P. Jesper Sjöström, and Thomas D. Mrsic-Flogel. Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473(7345):87–91, 2011.
- [9] Sonja B Hofer, Ho Ko, Bruno Pichler, Joshua Vogelstein, Hana Ros, Hongkui Zeng, Ed Lein, Nicholas A Lesica, and Thomas D Mrsic-Flogel. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neuroscience*, 14(8):1045–52, 2011.
- [10] Lee Cossell, Maria Florencia Iacaruso, Dylan R. Muir, Rachael Houlton, Elie N. Sader, Ho Ko, Sonja B. Hofer, and Thomas D. Mrsic-Flogel. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*, 518(7539):399–403, 2015.
- [11] Carsten K. Pfeffer, Mingshan Xue, Miao He, Z. Josh Huang, and Massimo Scanziani. Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons. *Nature Neuroscience*, 16(8):1068–1076, 2013.
- [12] Adil G. Khan, Jasper Poort, Angus Chadwick, Antonin Blot, Maneesh Sahani, Thomas D. Mrsic-Flogel, and Sonja B. Hofer. Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nature Neuroscience*, 21(6):851–859, 2018.
- [13] Hugh R Wilson and Jack D Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24, 1972.
- [14] Daniel J Amit and Nicolas Brunel. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 7(3):237–252, 1997.
- [15] Yashar Ahmadian, Daniel B. Rubin, and Kenneth D. Miller. Analysis of the Stabilized Supralinear Network. *Neural Computation*, 25(8):1994–2037, aug 2013.
- [16] C van Vreeswijk and H Sompolinsky. Chaotic balanced state in a model of cortical circuits. *Neural computation*, 10(6):1321–71, 1998.
- [17] Bilal Haider, Alvaro Duque, Andrea R Hasenstaub, and David A McCormick. Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *Journal of Neuroscience*, 26(17):4535–4545, 2006.
- [18] William R Softky and Christof Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *Journal of neuroscience*, 13(1):334–350, 1993.
- [19] Matteo Carandini and David J Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, 2012.
- [20] Yashar Ahmadian and Kenneth D. Miller. What is the dynamical regime of cerebral cortex? *Neuron*, 109(21):3373–3391, 2021.
- [21] Gianluigi Mongillo, David Hansel, and Carl Van Vreeswijk. Bistability and spatiotemporal irregularity in neuronal networks with nonlinear synaptic transmission. *Physical review letters*, 108(15):158101, 2012.
- [22] Ramin Khajeh, Francesco Fumarola, and LF Abbott. Sparse balance: excitatory-inhibitory networks with small bias currents and broadly distributed synaptic weights. *bioRxiv*, 2021.
- [23] Cody Baker, Vicky Zhu, and Robert Rosenbaum. Nonlinear stimulus representations in neural circuits with approximate excitatory-inhibitory balance. *PLOS Computational Biology*, 16(9):1–30, 09 2020.
- [24] Robert Rosenbaum and Brent Doiron. Balanced networks of spiking neurons with spatially dependent recurrent connections. *Physical Review X*, 4(2):021039, 2014.
- [25] Misha V Tsodyks, William E Skaggs, Terrence J Sejnowski, and Bruce L McNaughton. Paradoxical effects of external modulation of inhibitory interneurons. *Journal of neuroscience*, 17(11):4382–4388, 1997.
- [26] Daniel B Rubin, Stephen D Van Hooser, and Kenneth D Miller. The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*, 85(2):402–417, 2015.
- [27] Alessandro Sanzeni, Mark H Histed, and Nicolas Brunel. Response nonlinearities in networks of spiking neurons. *PLoS computational biology*, 16(9):e1008165, 2020.
- [28] Nataliya Kraynyukova and Tatjana Tchumatchenko. Stabilized supralinear network can give rise to bistable, oscillatory, and persistent activity. *Proceedings of the National Academy of Sciences*, 115(13):3464–3469, 2018.
- [29] Sadra Sadeh and Claudia Clopath. Inhibitory stabilization and cortical computation. *Nature Reviews Neuro-*

- science*, 22(1):21–37, January 2021.
- [30] Troy W Margrie, Michael Brecht, and Bert Sakmann. In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflügers Archiv*, 444(4):491–498, 2002.
- [31] Ede A Rancz, Kevin M Franks, Martin K Schwarz, Bruno Pichler, Andreas T Schaefer, and Troy W Margrie. Transfection via whole-cell recording in vivo: bridging single-cell physiology, genetics and connectomics. *Nature neuroscience*, 14(4):527–532, 2011.
- [32] Keith B Hengen, Mary E Lambo, Stephen D Van Hooser, Donald B Katz, and Gina G Turrigiano. Firing rate homeostasis in visual cortex of freely behaving rodents. *Neuron*, 80(2):335–342, 2013.
- [33] Luc J Gentet, Yves Kremer, Hiroki Taniguchi, Z Josh Huang, Jochen F Staiger, and Carl CH Petersen. Unique functional properties of somatostatin-expressing gabaergic neurons in mouse barrel cortex. *Nature neuroscience*, 15(4):607–612, 2012.
- [34] Daniel Jercog, Alex Roxin, Peter Bartho, Artur Luczak, Albert Compte, and Jaime de la Rocha. Up-down cortical dynamics reflect state transitions in a bistable network. *Elife*, 6:e22425, 2017.
- [35] Andrew YY Tan, Brandon D Brown, Benjamin Scholl, Deepankar Mohanty, and Nicholas J Priebe. Orientation selectivity of synaptic input to neurons in mouse and cat primary visual cortex. *Journal of Neuroscience*, 31(34):12339–12350, 2011.
- [36] Alex Roxin, Nicolas Brunel, David Hansel, Gianluigi Mongillo, and Carl van Vreeswijk. On the distribution of firing rates in networks of cortical neurons. *Journal of Neuroscience*, 31(45):16217–16226, 2011.
- [37] Bassam V Atallah, William Bruns, Matteo Carandini, and Massimo Scanziani. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron*, 73(1):159–170, 2012.
- [38] Wulfram Gerstner and Werner M Kistler. *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press, 2002.
- [39] Luigi M Ricciardi. The first passage time problem. In *Diffusion Processes and Related Topics in Biology*, pages 61–72. Springer, 1977.
- [40] Nicholas J Priebe, Ferenc Mechler, Matteo Carandini, and David Ferster. The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nature Neuroscience*, 7(10):1113–1122, 2004.
- [41] A Destexhe, M Rudolph, J M Fellous, and T J Sejnowski. Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience*, 107(1):13–24, 2001.
- [42] Tatjana Tchumatchenko, Aleksey Malyshev, Theo Geisel, Maxim Volgushev, and Fred Wolf. Correlations and synchrony in threshold neuron models. *Phys. Rev. Lett.*, 104:058102, Feb 2010.
- [43] Jason M Guest, Arco Bast, Rajeevan T Narayanan, and Marcel Oberlaender. Thalamus gates active dendritic computations in cortex during sensory processing. *bioRxiv*, 2021.
- [44] Sean E Cavanagh, Laurence T Hunt, and Steven W Kennerley. A diversity of intrinsic timescales underlie neural computations. *Frontiers in Neural Circuits*, 14:81, 2020.
- [45] Sean E Cavanagh, Joni D Wallis, Steven W Kennerley, and Laurence T Hunt. Autocorrelation structure at rest predicts value correlates of single neurons during reward-guided choice. *elife*, 5:e18937, 2016.
- [46] Guozhang Chen and Pulin Gong. Computing by modulating spontaneous cortical activity patterns as a mechanism of active visual processing. *Nature communications*, 10(1):1–15, 2019.
- [47] Kenneth D Miller and Agostina Palmigiano. Generalized paradoxical effects in excitatory/inhibitory networks. *bioRxiv*, 2020.
- [48] Yue Kris Wu and Friedemann Zenke. Nonlinear transient amplification in recurrent neural networks with short-term plasticity. *Elife*, 10:e71263, 2021.
- [49] Laura Busse, Alex R. Wade, and Matteo Carandini. Representation of Concurrent Stimuli by Population Activity in Visual Cortex. *Neuron*, 64(6):931–942, 2009.
- [50] Dario Ringach and Robert Shapley. Reverse correlation in neurophysiology. *Cognitive Science*, 28(2):147–166, 2004.
- [51] Denise M Piscopo, Rana N El-Danaf, Andrew D Huberman, and Christopher M Niell. Diverse visual features encoded in mouse lateral geniculate nucleus. *Journal of Neuroscience*, 33(11):4642–4656, 2013.
- [52] Erez Persi, David Hansel, Lionel Nowak, Pascal Barone, and Carl Van Vreeswijk. Power-law input-output transfer functions explain the contrast-response and tuning properties of neurons in visual cortex. *PLoS Comput Biol*, 7(2):e1001078, 2011.
- [53] Davi D. Bock, Wei Chung Allen Lee, Aaron M. Kerlin, Mark L. Andermann, Greg Hood, Arthur W. Wetzel, Sergey Yurgenson, Edward R. Soucy, Hyon Suk Kim, and R. Clay Reid. Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471(7337):177–184, mar 2011.
- [54] Stephanie C Seeman, Luke Campagnola, Pasha A Davoudian, Alex Hoggarth, Travis A Hage, Alice Bosma-Moody, Christopher A Baker, Jung Hoon Lee, Stefan Mihalas, Corinne Teeter, et al. Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. *Elife*, 7:e37349, 2018.
- [55] John D Murray, Alberto Bernacchia, David J Freedman, Ranulfo Romo, Jonathan D Wallis, Xinying Cai, Camillo Padoa-Schioppa, Tatiana Pasternak, Hyojung Seo, Daeyeol Lee, et al. A hierarchy of intrinsic timescales across primate cortex. *Nature neuroscience*, 17(12):1661–1663, 2014.
- [56] Marvin Minsky and Seymour Papert. Perceptrons. 1969.
- [57] Hirofumi Ozeki, Ian M Finn, Evan S Schaffer, Kenneth D Miller, and David Ferster. Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62(4):578–592, 2009.
- [58] Alessandro Sanzeni, Bradley Akitake, Hannah C Goldbach, Caitlin E Leedy, Nicolas Brunel, and Mark H Histed. Inhibition stabilization is a widespread property of cortical networks. *eLife*, 9:e54875, June 2020.
- [59] Joaquin M Fuster and Garrett E Alexander. Neuron activity related to short-term memory. *Science*, 173(3997):652–654, 1971.
- [60] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Matteo Carandini, and Kenneth D Harris. High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365, 2019.
- [61] Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437), 2019.

- [62] Michael Okun and Ilan Lampl. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature neuroscience*, 11(5):535–537, 2008.
- [63] Peter Rupprecht and Rainer W Friedrich. Precise synaptic balance in the zebrafish homolog of olfactory cortex. *Neuron*, 100(3):669–683, 2018.
- [64] Anthony D Lien and Massimo Scanziani. Tuned thalamic excitation is amplified by visual cortical circuits. *Nature Neuroscience*, 16(9):1315–1323, 2013.
- [65] Cindy Poo and Jeffry S Isaacson. A major role for intracortical circuits in the strength and tuning of odor-evoked excitation in olfactory cortex. *Neuron*, 72(1):41–48, 2011.
- [66] Ya-tang Li, Leena A Ibrahim, Bao-hua Liu, Li I Zhang, and Huizhong Whit Tao. Linear transformation of thalamocortical input by intracortical excitation. *Nature neuroscience*, 16(9):1324–1330, 2013.
- [67] James H Marshel, Yoon Seok Kim, Timothy A Machado, Sean Quirin, Brandon Benson, Jonathan Kadmon, Cephra Raja, Adelaida Chibukhchyan, Charu Ramakrishnan, Masatoshi Inoue, et al. Cortical layer-specific critical dynamics triggering perception. *Science*, 365(6453), 2019.
- [68] Sandrine Lefort, Christian Tamm, J-C Floyd Sarria, and Carl CH Petersen. The excitatory neuronal network of the c2 barrel column in mouse primary somatosensory cortex. *Neuron*, 61(2):301–316, 2009.
- [69] Marcel Oberlaender, Christiaan PJ de Kock, Randy M Bruno, Alejandro Ramirez, Hanno S Meyer, Vincent J Dercksen, Moritz Helmstaedter, and Bert Sakmann. Cell type-specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissa cortex. *Cerebral cortex*, 22(10):2375–2391, 2012.
- [70] Christiaan PJ de Kock, Jean Pie, Anton W Pieneman, Rebecca A Mease, Arco Bast, Jason M Guest, Marcel Oberlaender, Huibert D Mansvelder, and Bert Sakmann. High-frequency burst spiking in layer 5 thick-tufted pyramids of rat primary somatosensory cortex encodes exploratory touch. *Communications biology*, 4(1):1–14, 2021.
- [71] Joe W Graham, Michael Emiel Gevaert, Thomas K Berger, Guy Antoine Atenekeng Kahou, Carlos Aguado Sanchez, Michael W Reimann, Henry Markram, Eilif Muller, Srikanth Ramaswamy, Marwan Abdellah, et al. Reconstruction and simulation of neocortical microcircuitry. *Cell*, 163(2), 2015.
- [72] Michael Stobb, Joshua M Peterson, Borbala Mazzag, and Ethan Gahtan. Graph theoretical model of a sensorimotor connectome in zebrafish. *PLoS One*, 7(5):e37292, 2012.
- [73] Yazan N. Billeh, Binghuang Cai, Sergey L. Gratiy, Kael Dai, Ramakrishnan Iyer, Nathan W. Gouwens, Reza Abbasi-Asl, Xiaoxuan Jia, Joshua H. Siegle, Shawn R. Olsen, Christof Koch, Stefan Mihalas, and Anton Arkhipov. Systematic Integration of Structural and Functional Data into Multi-scale Models of Mouse Primary Visual Cortex. *Neuron*, 106(3):388–403.e18, May 2020.
- [74] Lukasz Kuśmiercz, Shun Ogawa, and Taro Toyozumi. Edge of chaos and avalanches in neural networks with heavy-tailed synaptic weight distribution. *Physical Review Letters*, 125(2):028101, 2020.
- [75] Suzana Herculano-Houzel, Charles R Watson, and George Paxinos. Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. *Frontiers in neuroanatomy*, 7:35, 2013.
- [76] Daniel J Amit and MV Tsodyks. Quantitative study of attractor neural network retrieving at low spike rates. i. substrate-spikes, rates and neuronal gain. *Network: Computation in neural systems*, 2(3):259, 1991.
- [77] Sophia Becker, Andreas Nold, and Tatjana Tchumachenko. Formation and synaptic control of active transient working memory representations. *BioRxiv*, 2020.
- [78] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208, 2000.
- [79] Carl CH Petersen. Whole-cell recording of neuronal membrane potential during behavior. *Neuron*, 95(6):1266–1281, 2017.
- [80] Anton Arkhipov, Nathan W Gouwens, Yazan N Billeh, Sergey Gratiy, Ramakrishnan Iyer, Ziqiang Wei, Zihao Xu, Reza Abbasi-Asl, Jim Berg, Michael Buice, et al. Visual physiology of the layer 4 cortical circuit in silico. *PLoS Computational Biology*, 14(11):e1006535, 2018.
- [81] Stephen D Van Hooser, J Alexander F Heimel, Sooyoung Chung, Sacha B Nelson, and Louis J Toth. Orientation selectivity without orientation maps in visual cortex of a highly visual mammal. *Journal of Neuroscience*, 25(1):19–28, 2005.
- [82] DH Hubel, TN Wiesel, and S LeVay. Functional architecture of area 17 in normal and monocularly deprived macaque monkeys. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 40, pages 581–589. Cold Spring Harbor Laboratory Press, 1976.
- [83] Klaus Obermayer and Gary G Blasdel. Geometry of orientation and ocular dominance columns in monkey striate cortex. *Journal of Neuroscience*, 13(10):4114–4129, 1993.
- [84] Moshe Gur and D Max Snodderly. Physiological differences between neurons in layer 2 and layer 3 of primary visual cortex (v1) of alert macaque monkeys. *The Journal of physiology*, 586(9):2293–2306, 2008.
- [85] John O’Kusky and Marc Colonnier. A laminar analysis of the number of neurons, glia, and synapses in the visual cortex (area 17) of adult macaque monkeys. *Journal of Comparative Neurology*, 210(3):278–290, 1982.
- [86] Srikanth Ramaswamy, Jean-Denis Courcol, Marwan Abdellah, Stanislaw R Adaszewski, Nicolas Antille, Selim Arsever, Guy Atenekeng, Ahmet Bilgili, Yury Brukau, Athanassia Chalimourda, et al. The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Frontiers in neural circuits*, 9:44, 2015.
- [87] Daniel Keller, Csaba Erő, and Henry Markram. Cell densities in the mouse brain: a systematic review. *Frontiers in neuroanatomy*, 12:83, 2018.
- [88] Philip L Smith. From poisson shot noise to the integrated ornstein–uhlenbeck process: Neurally principled models of information accumulation in decision-making and response time. *Journal of Mathematical Psychology*, 54(2):266–283, 2010.
- [89] Sheldon M Ross, John J Kelly, Roger J Sullivan, William James Perry, Donald Mercer, Ruth M Davis, Thomas Dell Washburn, Earl V Sager, Joseph B Boyce, and Vincent L Bristow. *Stochastic processes*, volume 2. Wiley New York, 1996.