

# Republished: Dynamics of Stochastic Integrate-and-Fire Networks

Gabriel Koch Ocker<sup>\*</sup>

*Department of Mathematics and Statistics, Boston University, Boston, Massachusetts 02215, USA*

 (Received 19 April 2023; accepted 7 November 2023; published 11 December 2023)

The neural dynamics generating sensory, motor, and cognitive functions are commonly understood through field theories for neural population activity. Classic neural field theories are derived from highly simplified models of individual neurons, while biological neurons are highly complex cells. Integrate-and-fire models retain a key nonlinear feature of neuronal activity: Action potentials return the membrane potential to a nearly fixed reset value. This nonlinear reset of the membrane voltage after a spike is absent from classic neural field theories. Here, we develop a statistical field theory for networks of integrate-and-fire neurons with stochastic spike emission. This reveals a new mean-field theory for the activity in these networks, fluctuation corrections to the mean-field dynamics, and a mapping to a self-consistent renewal process. We use these to study the impact of the spike-driven reset of the membrane voltage on population activity. The spike reset gives rise to a multiplicative, rate-dependent leak term in the mean-field membrane voltage dynamics. This leads to bistability between quiescent and active states in the mean-field theory of homogeneous and excitatory-inhibitory pulse-coupled networks. We uncover two types of fluctuation correction to the mean-field theory, due to the nonlinear mapping from membrane voltage to spike emission and the nonlinear reset. These can have competing effects, promoting and suppressing activity, respectively. We then examine the roles of spike resets and recurrent inhibition in stabilizing network activity. We calculate the phase diagram for inhibitory stabilization and find that an inhibition-stabilized regime occurs in wide regions of parameter space, consistent with experimental reports of inhibitory stabilization in diverse brain regions. Fluctuations narrow the region of inhibitory stabilization, consistent with their role in suppressing activity through spike resets.

DOI: [10.1103/PhysRevX.13.041047](https://doi.org/10.1103/PhysRevX.13.041047)

Subject Areas: Biological Physics, Statistical Physics

## I. INTRODUCTION

The activity of neuronal populations underlies sensory, motor, and cognitive functions. Mathematical theories for predicting the macroscopic activity of neural populations are a core tool of computational neuroscience, psychology, and psychiatry [1–4]. These theories typically rely on neural activity equations, with variants also called rate equations, neural mass equations, or, if placed on a spatial domain, neural field equations:

$$\partial_t \mathbf{v} = -\mathbf{v} + \mathbf{E} + \mathbf{J} * \phi(\mathbf{v}), \quad (1)$$

where bold terms denote a vector or matrix-valued function,  $\phi$  is a single-unit nonlinearity applied elementwise,  $\partial_t$  is the derivative with respect to time, and  $*$  is a

matrix convolution:  $\mathbf{J} * \phi(\mathbf{v}) = \sum_j \int ds J_{ij}(s) \phi(v_j(t-s))$ . These and similar equations are commonly understood as a coarse-grained model for large populations of neurons [5–9]. Formally, they are a mean-field theory for populations of neurons that switch between discrete active and quiescent states [10–13], or for generalized linear point process models [Eq. (9)].

Biological neurons' membrane voltages, however, have complex nonlinear dynamics [14]. Neural field equations have been supplemented with some biophysical detail in an *ad hoc* fashion [3]. A principled mean-field theory of more biophysical neuron models would expose how single-neuron biophysics shape macroscopic population activity [15].

Integrate-and-fire models, which replace the nonlinear dynamics of spike generation by a simple fire-and-reset rule for the membrane voltage, are fruitful tools for investigating how network structure and synaptic and neuronal biophysics shape macroscopic activity [16,17]. The classic mean-field theory of integrate-and-fire networks focuses on the density of membrane voltages across a population [18]. If the net recurrent input to each neuron is a white Gaussian process, the membrane voltage density obeys a Fokker-Planck partial differential equation [19]. Numerical or

<sup>\*</sup>gkocker@bu.edu

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

special function solutions of that Fokker-Planck equation expose steady-state and weakly nonequilibrium population firing rates and pairwise statistics [20–24].

The assumption of white Gaussian input currents is, however, inconsistent with the resulting temporally colored spike train statistics [25]. In some cases, the Fokker-Planck approach for the population voltage density can be extended to temporally structured fluctuations [26–28]. Alternatively, for generalized integrate-and-fire neurons with stochastic spike emission, population firing rates and pairwise statistics can be predicted from the density of interspike times rather than the density of membrane voltages [29–34]. Population density approaches expose approximate low-dimensional dynamics through eigenfunctions of the density evolution operator [35,36].

Here, we study the integrate-and-fire model with stochastic spike emission. We construct the full joint probability density functional of a neuronal network’s spike trains and membrane voltages using the response variable path integral formalism [37–40]. This formalism is commonly applied to nonspiking models, where it has exposed chaotic and metastable regimes [41–44], the memory capacity of recurrent networks [45,46], and the dynamical impact of computationally or biologically constrained connectivity [47–50]. It has also been applied to spiking models without spike resets or in a phase formulation [51–58].

The joint density functional exposes a new simple, deterministic mean-field theory for stochastic integrate-and-fire networks: activity equations like Eq. (1) with an additional rate-dependent leak. This novel nonlinearity qualitatively shapes networks’ macroscopic dynamics. We study networks in an increasing order of complexity, progressing from uncoupled neurons to single-population recurrent networks and then networks with multiple cell types. Spike resets can stabilize strongly coupled excitatory networks with unbounded spike intensity functions. We uncover bistable regimes in homogeneous and excitatory-inhibitory networks. Examining the impact of fluctuations on the activity using renewal theory and a self-consistent Gaussian approximation with colored noise, we find that due to the nonlinearity of the spike reset, fluctuations suppress activity.

In the classic neural activity equations, inhibitory feedback is necessary to stabilize strong recurrent excitation [8,59]. A paradoxical reduction of inhibitory activity after inhibitory stimulation is a signature of an inhibition-stabilized regime [60] and is observed in diverse mammalian cortices [61–64]. We find that the phase diagram for excitatory-inhibitory networks includes wide regions of paradoxical responses, suggesting a generic mechanism for their widespread experimental observation. Spiking fluctuations narrow the region of inhibitory stabilization, consistent with their intrinsically stabilizing effect through resets of the membrane voltage.

## II. STOCHASTIC INTEGRATE-AND-FIRE MODEL

We introduce the stochastic leaky integrate-and-fire (LIF) model in discrete time first and then take a continuous-time limit. At each small time step  $t \in [T]$ , of width  $dt$ , neuron  $i \in [N]$  generates  $dn_{it} \in \{0, 1\}$  spikes. ( $n_{it}$  is the cumulative spike count of neuron  $i$  at time  $t$ .) Neuron  $i$  receives inputs  $d\mathbf{n}$  through weighted synaptic filters  $\mathbf{J}$ . It also has a resting voltage  $E_i$ , which may depend on external applied currents. We take  $dn_{it}$  to be generated as a Bernoulli random variable with spike probability  $f(v_i)dt$ , for some intensity function  $0 \leq f(v) \leq dt^{-1}$ . After a spike is emitted, that neuron’s membrane voltage is reset to within  $\mathcal{O}(dt)$  of the reset value  $r$ . If  $f(v) = \theta(v - b)/dt$ , where  $\theta(x)$  is the Heaviside step function, the deterministic LIF neuron with threshold  $b$  is recovered [65]. In the continuous-time limit (Appendix A),

$$\partial_t \mathbf{v}(t) = \frac{1}{\tau} [-\mathbf{v}(t) + \mathbf{E}(t) + (\mathbf{J} * \dot{\mathbf{n}})(t)] - \dot{\mathbf{n}}(t^+) [\mathbf{v}(t) - r]. \quad (2)$$

Here,  $\dot{n}_i(t^+) \equiv \partial_t n_i(t^+)$ , and  $t^+ = t + \epsilon$  for an infinitesimal  $\epsilon > 0$  so the membrane potential is reset at time  $t$  by an immediately preceding spike. Each  $\dot{n}_i(t)$  is an inhomogeneous Poisson process with intensity  $f(v_i(t))$ . The Poisson spike emission arises as the continuous-time limit of the discrete-time Bernoulli spike train. The last term in Eq. (2) is the reset of the membrane voltage after a spike. This nonlinear coupling between the spike train and membrane voltage is the key feature of this model compared to generalized linear models. (See Appendix B for a discussion of absolute refractory periods in this model.) We will nondimensionalize the model, measuring time relative to  $\tau$  and shifting  $\mathbf{v}$  and  $\mathbf{E}$  by the reset  $r$ .

Equation (2) is a set of coupled stochastic differential equations with multiplicative Poisson noise. The expected trajectory obeys

$$\begin{aligned} \partial_t \langle \mathbf{v} \rangle &= -\langle \mathbf{v} \rangle + \mathbf{E} + \mathbf{J} * \langle \dot{\mathbf{n}} \rangle - \langle \dot{\mathbf{n}} \rangle \langle \mathbf{v} \rangle - \langle \langle \dot{\mathbf{n}} \mathbf{v} \rangle \rangle, \\ \langle \dot{\mathbf{n}} \rangle &= \langle f(\mathbf{v}) \rangle, \end{aligned} \quad (3)$$

where  $\langle \rangle$  denotes a moment and  $\langle \langle \rangle \rangle$  denotes a cumulant. (Here we suppress the explicit time dependencies, as well as the infinitesimal time shift in the reset term. Moving forward, we will often continue to suppress those.) To calculate those cumulants requires the joint density functional of the membrane voltages and spike trains. In the response variable path integral formalism, it is (Appendix A)

$$\begin{aligned} p[\mathbf{v}, \dot{\mathbf{n}}] &= \int \mathcal{D}\tilde{\mathbf{v}} \int \mathcal{D}\tilde{\mathbf{n}} \exp(-S[\mathbf{v}, \dot{\mathbf{n}}, \tilde{\mathbf{v}}, \tilde{\mathbf{n}}]), \\ S[\mathbf{v}, \dot{\mathbf{n}}, \tilde{\mathbf{v}}, \tilde{\mathbf{n}}] &= \tilde{\mathbf{v}}^T (\partial_t \mathbf{v} + \mathbf{v} - \mathbf{E} - \mathbf{J} * \dot{\mathbf{n}} + \dot{\mathbf{n}} \mathbf{v}) \\ &\quad + \tilde{\mathbf{n}}^T \dot{\mathbf{n}} - [\exp(\tilde{\mathbf{n}}) - 1]^T \mathbf{f}. \end{aligned} \quad (4)$$

Here,  $\mathbf{x}^T \mathbf{y} = \sum_i \int dt x_i(t) y_i(t)$  is the functional inner product and  $f_i(t) = f(v_i(t))$ .  $S$  is the action functional.  $\tilde{\mathbf{n}}, \tilde{\mathbf{v}}$  are purely imaginary auxiliary variables, called the response variables because joint moments with them measure responses to fluctuations in the activity.

This density has the  $N$ -dimensional deterministic mean-field theory:

$$\partial_t \tilde{\mathbf{v}} = -\tilde{\mathbf{v}} + \mathbf{E} + \mathbf{J} * \tilde{\mathbf{f}} - \tilde{\mathbf{f}} \tilde{\mathbf{v}}. \quad (5)$$

The mean-field value of  $\dot{n}_i$  is  $\tilde{f}_i = f(\tilde{v}_i)$ . The  $N$ -dimensional mean-field theory is an approximation of Eq. (3). If we assume that

- (1) the expectation of the spike trains is  $\langle \dot{\mathbf{n}} \rangle = f(\langle \mathbf{v} \rangle)$
- (2) the spikes and membrane voltage are independent so that the joint cumulant  $\langle\langle \dot{\mathbf{n}} \mathbf{v} \rangle\rangle = 0$ ,

then Eq. (3) reduces to Eq. (5). Assumption (1) is correct only if  $f$  is linear. Assumption (2) is generally incorrect, although it may be a good approximation if  $\langle \mathbf{v} \rangle \langle \dot{\mathbf{n}} \rangle \gg \langle\langle \dot{\mathbf{n}} \mathbf{v} \rangle\rangle$ .

Formally, we expand the configuration variables  $\mathbf{v}, \dot{\mathbf{n}}$  around their mean values. The mean-field theory is then the result of a saddle point approximation for integrals over the fluctuations (Appendix E). This corresponds to assuming fluctuations are negligible so  $p[\dot{\mathbf{n}}] = \delta[\dot{\mathbf{n}} - f(\mathbf{v})]$ , or equivalently, truncating the action at linear order in  $\tilde{\mathbf{n}}$ . This implies assumptions (1) and (2), so Eq. (3) reduces to Eq. (5). This approach also exposes fluctuation corrections to the mean-field theory. We will see in the next section that the two nonlinearities in Eq. (2) impart different fluctuation corrections to the mean-field theory. First, we compare the mean-field theory, Eq. (5), to the classic activity equations, Eq. (1).

The mean-field dynamics of Eq. (5) differ from Eq. (1) in two ways. The first is the presence of the reset term  $-f(v_i)v_i$ . The second is in the interpretation of the nonlinearity  $f$ . Here,  $f$  determines the instantaneous spike emission probability as a function of the membrane voltage and is typically required to be nonsaturating so that the neuron is guaranteed to spike if  $v_i \rightarrow \infty$ . (This is not mathematically necessary;  $f$  could be chosen to saturate at a finite value. In discrete time,  $f$  must be bounded by  $1/dt$  so the spike probability does not exceed 1.) In the microscopic binary switching model underlying Eq. (1), the nonlinearity  $\phi$  determines the single-neuron transition rates from quiescence to activity and is typically chosen as a sigmoid to prevent unbounded activity. In either case, the nonlinearity  $f$  or  $\phi$  is a property of individual neurons.

Can we map the new mean-field theory, Eq. (5), onto the classic activity equations, Eq. (1), with an effective nonlinearity  $\phi$  that includes the effect of the rate-dependent leak? Requiring  $\mathbf{J}\phi = \mathbf{J}\mathbf{f} - \mathbf{v}\mathbf{f}$ , with  $(\mathbf{v}\mathbf{f})_i = v_i f(v_i)$ , we find that if the coupling  $\mathbf{J}$  has a left inverse,

$$\phi(\mathbf{v}) = \mathbf{f} - \mathbf{J}^{-1}(\mathbf{v}\mathbf{f}). \quad (6)$$

So to map the mean-field theory of Eq. (5) onto the classic activity equations, the effective nonlinearity  $\phi$  depends explicitly on the coupling  $\mathbf{J}$ ; it is no longer a single-neuron nonlinearity. If there are linear self-interactions and inter-neuronal coupling is weak so that  $\mathbf{J}$  is diagonally dominant, the effective nonlinearity will be approximately a single-neuron property.

The other classic form of rate equation is  $\tau \partial_t v = -v + \phi(\mathbf{J} * \mathbf{v} + \mathbf{E})$ . This is also a mean-field theory of binary switching neurons [11–13]. Here,  $v$  is commonly understood as a mean-field description of the firing rate or proportion of active neurons in a population, rather than the membrane potential or synaptic drive [8,9]. The two types of activity equation differ in their assumptions about the dominant synaptic or neuronal timescales [13,66].

To map Eq. (5) onto this classic activity equation would require  $\phi((\mathbf{J} * \mathbf{v})_i + E_i) = -v_i f(v_i) + \sum_j J_{ij} * f(v_j) + E_i$ . In general, to map Eq. (5) onto the classic activity equations may require the nonlinearity to be a function of the coupling operator, activity variable, and baseline drive separately, rather than a function of their sum.

Mapping Eq. (5) onto the classic activity equations can thus introduce nonlinearities tailored to a specific LIF network, rather than as single-neuron input-rate functions. This mapping is, however, not necessary. The mean-field dynamics of Eq. (5) are amenable to direct analysis.

### III. IMPACT OF SPIKE RESET AND FLUCTUATIONS ON SINGLE-NEURON ACTIVITY

We now examine the steady-state input-rate transfer of a single neuron or, equivalently, an uncoupled population. The mean-field firing rate  $\tilde{f}$  is given by equilibria of Eq. (5) with  $\mathbf{J} = \mathbf{0}$  and constant  $E$ . We consider neurons with threshold-power law spike probability functions,  $f(v) = [v - 1]_+^a$ , which match the effective nonlinearity of mechanistic spiking models and biological neurons in fluctuation-driven regimes [67–71]. (The membrane voltage has been nondimensionalized to set the threshold for spike generation at  $v = 1$ .) For simplicity, we take a threshold-linear neuron with  $f(v) = [v - 1]_+$  so the equilibrium solution to the mean-field equation is

$$\tilde{f} = [\sqrt{E} - 1]_+. \quad (7)$$

The mean-field theory for the stochastic LIF neuron predicts its equilibrium firing rate as a function of its membrane voltage [Fig. 1(b), black line vs dots]. At higher rates, Eq. (7) overpredicts the true firing rates. Since the mean-field theory neglects all fluctuations, fluctuations suppress activity in the stochastic LIF model.

For comparison, consider a stochastic LIF model with a linear reset: Each spike causes a decrease in the membrane voltage of size  $r$  [Fig. 1(a), blue line] [72]. The action for that model is

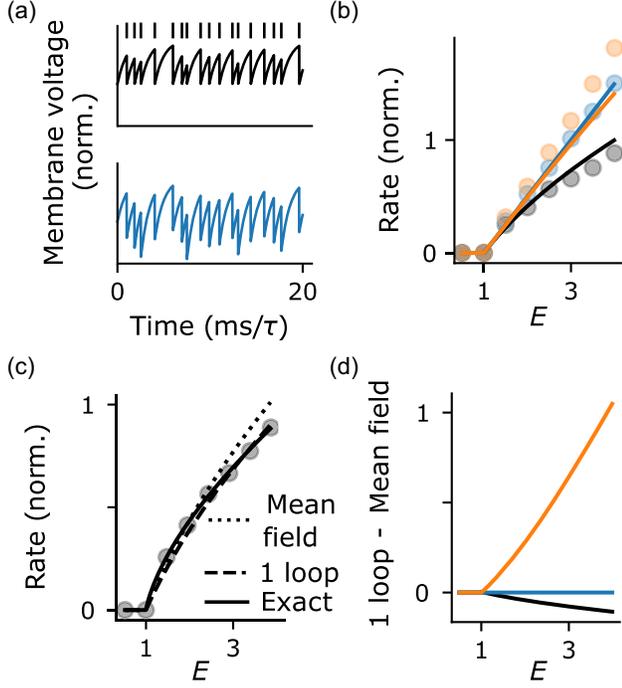


FIG. 1. Impact of fluctuations on firing rates through spike resets and nonlinear intensity functions. (a) Membrane voltage traces of the stochastic LIF neuron (top, black) and a neuron with linear resets (bottom, blue). For comparison, the two neurons are forced to have the same spike times (in this panel only). (b) Firing rate vs resting voltage  $E$  for three models. Black: the stochastic LIF neuron with a threshold-linear intensity function,  $f(v) = [v - 1]_+$ . Blue: the linear-reset model with a matched intensity function,  $f_L(v) = f(v)$ . Orange: the linear-reset model with matched mean-field membrane voltages,  $f_M(v) = vf(v)$ . Dots: simulation. Solid curves: mean-field predictions [Eqs. (7), (10), and (11)]. (c) Impact of fluctuations on the stochastic LIF neuron's firing rate. Dotted line: mean-field prediction. Dashed line: self-consistent one-loop prediction, accounting for Gaussian fluctuations around the expected voltage and rate [Eq. (12)]. Solid line: exact renewal theory prediction from Eq. (17). (d) Difference between the one-loop and mean-field rates for the three models.

$$S[\mathbf{v}, \dot{\mathbf{n}}, \tilde{\mathbf{v}}, \tilde{\mathbf{n}}] = \tilde{\mathbf{v}}^T (\partial_t \mathbf{v} + \mathbf{v} - \mathbf{E} - \mathbf{J} * \dot{\mathbf{n}} + \dot{\mathbf{n}} r) + \tilde{\mathbf{n}}^T \dot{\mathbf{n}} - [\exp(\tilde{\mathbf{n}}) - 1]^T \mathbf{f}_L, \quad (8)$$

with the  $N$ -dimensional mean-field theory:

$$\partial_t \bar{\mathbf{v}} = -\bar{\mathbf{v}} + \mathbf{E} + \mathbf{J} * \bar{\mathbf{f}} - \bar{\mathbf{f}} r. \quad (9)$$

This has a similar form to the classic activity equation, Eq. (1), and can be directly mapped onto it with the substitution  $J_{ii}(s) \rightarrow J_{ii}(s) - r\delta(s)$ . For this reason, we say that Eq. (1) is a mean-field theory for a stochastic LIF neuron with linear resets, which is an example of a generalized linear model or zeroth-order spike response model [73]. The mean-field firing rate of the uncoupled linear-reset model, with  $r = v_{\text{th}} = 1$ , is

$$\bar{f}_L = \left[ \frac{E - 1}{2} \right]_+. \quad (10)$$

For a perithreshold stimulus,  $E = 1 + \epsilon$  in Eqs. (7) and (10),  $\bar{f} = \epsilon/2 + \mathcal{O}(\epsilon^2) \approx \bar{f}_L$  and the mean-field theories of the stochastic LIF and linear-reset models match for infinitesimal firing rates. At finite rates, however, the linear-reset model provides a poor prediction for the stochastic LIF neuron [Fig. 1(b), blue dots vs black dots].

Instead of matching the intensity functions of the two models, we could match their mean-field membrane voltage by giving the linear-reset model the intensity function  $f_M(v) = vf(v)$ . [ $f(v)$  is the intensity function of the stochastic LIF neuron.] The mean-field rate of this matched linear-reset model is

$$\bar{f}_M = \sqrt{E} [\sqrt{E} - 1]_+. \quad (11)$$

For the matched linear-reset model, the mean-field firing rate underpredicts the true activity level [Fig. 1(b), orange line vs dots], so fluctuations promote activity. Why do fluctuations suppress activity in the stochastic LIF model but promote activity in the matched linear-reset model?

In  $\partial_t \langle \mathbf{v} \rangle$ , we need to account for (1) the nonlinearity in the intensity function and (2) the nonlinear spike reset. To that end, we expand the membrane voltage and spike trains around their means to derive an expansion for the action that self-consistently accounts for the impact of fluctuations (the loop expansion of the effective action; see Appendix E). This allows us to derive diagrammatic corrections to the mean-field theory. Loop diagrams measure the influence of higher-order activity statistics on lower-order statistics. There are loop corrections to the mean-field theory when the model has a nonlinearity.

The stochastic LIF has two nonlinearities: the intensity function and the nonlinear spike reset. These give rise to fluctuation corrections in the mean voltage and rate:

$$0 = \partial_t \bar{v} + \bar{v} + \bar{v} \bar{n} - E + \text{loop diagram}, \quad (12)$$

$$0 = \bar{n} - f(\bar{v}) - \text{loop diagram}.$$

Without the one-loop diagrams, these reduce to the mean-field theory of Eq. (5) with  $\mathbf{J} = 0$ . The one-loop diagrams measure the impact of two-point fluctuations on the mean through the two nonlinearities of the intensity function and spike reset.

In field-theoretic terms, these loop diagrams represent proper vertex corrections to the effective action (Appendix E). We can also understand them by comparing Eq. (3) and Eq. (12). In their first lines, the loop diagram



is an approximation of  $\langle\langle \dot{n}v \rangle\rangle$ . In the second line of

Eq. (12), the loop diagram originates in a Taylor expansion of  $f(v)$  around  $f(\langle v \rangle)$  in Eq. (3); the loop diagram in the second line approximates  $(f^{(2)}/2)\langle\langle v^2 \rangle\rangle$ . We next discuss these approximations.

The edges in the Feynman diagrams correspond to factors of the linear response of the configuration to a fluctuation  $\bar{\Delta}$ , also called a propagator. Since the model has two configuration variables, each with a corresponding response variable, there are four types of propagator:

- (1) the spike response to a spike fluctuation  $\bar{\Delta}_{n,\bar{n}}$ ,
- (2) the voltage response to a spike fluctuation  $\bar{\Delta}_{v,\bar{n}}$ ,
- (3) the spike response to a voltage fluctuation  $\bar{\Delta}_{n,\bar{v}}$ , and
- (4) the voltage response to a voltage fluctuation  $\bar{\Delta}_{v,\bar{v}}$ .

We represent them with the edges

$$\begin{aligned}
 \bar{\Delta}_{n,\bar{n}} &= \text{———} \\
 \bar{\Delta}_{v,\bar{n}} &= \text{~~~~~} \\
 \bar{\Delta}_{n,\bar{v}} &= \text{~~~~~} \\
 \bar{\Delta}_{v,\bar{v}} &= \text{-----}
 \end{aligned} \tag{13}$$

Only the first two of these edges appear in the one-loop equations of motion for the mean voltage and rate. The vertex  $\bullet$  represents the intensity  $f(\bar{v})$ . Each diagram also has a vertex  $\circ$ . These vertices have different origins in the two diagrams, corresponding to either the spike reset or intensity function. (The two types of  $\circ$  vertex can be distinguished by their incoming edges.) The definition of the linear response functions for the stochastic LIF model are given in Appendix D 1, along with the Feynman rules for perturbative corrections to the mean-field theory (see also Appendix E).

In Eq. (3), at a stationary state,

$$\begin{aligned}
 &\langle\langle \dot{n}(t^+) v(t) \rangle\rangle \\
 &\approx \text{Diagram: a circle with a scalloped edge, a central dot, and a vertex } \circ \text{ on the left edge.} \\
 &= \lim_{t^+ \rightarrow t} \int \frac{d\omega}{2\pi} e^{i\omega(t^+ - t)} \bar{\Delta}_{n,\bar{n}}(\omega) \bar{\Delta}_{v,\bar{n}}(-\omega) f(\bar{v}) \\
 &= \frac{f^{(1)}\bar{v}^2}{2(1 + \bar{n} + f^{(1)}\bar{v})} f(\bar{v}).
 \end{aligned} \tag{14}$$

We have written the cumulant for a single neuron, dropping the neuron index implicit in Eq. (3), and leveraged the stationarity assumption so that  $\bar{\Delta}(t, t') = \bar{\Delta}(t' - t) = \mathcal{F}^{-1}[\bar{\Delta}(\omega)]$ , with the Fourier transform convention  $\mathcal{F}[\bar{\Delta}(s)] = \int ds \bar{\Delta}(s) e^{-i\omega s}$ . The limit  $t^+ \rightarrow t$  is taken from the right,  $t^+ > t$ .  $f^{(p)}$  is the order- $p$  derivative of the intensity function  $f$ , evaluated at  $\bar{v}$ . This diagram, and the vertex  $\circ$  in it, arise from the spike reset term of Eq. (2).

If the neuron is in an active steady state,  $\bar{v} > 1$  and  $f(\bar{v}), \bar{n} > 0$ . If  $f^{(1)} > 0$ , every term in this approximate cumulant is positive. It thus decreases  $\partial_t \bar{v}$  in Eq. (12).

The equation of motion governing  $\bar{n}$  also has a loop correction. In the second line of Eq. (3), we expand  $f(v)$  around the mean. Truncating at second order, in a stationary state

$$\begin{aligned}
 \frac{f^{(2)}}{2} \langle\langle v(t) v(t) \rangle\rangle &\approx \text{Diagram: a circle with a scalloped edge and a central dot.} \\
 &= \frac{f^{(2)}}{2} \int \frac{d\omega}{2\pi} \bar{\Delta}_{v,\bar{n}}(\omega) \bar{\Delta}_{v,\bar{n}}(-\omega) f(\bar{v}) \\
 &= \frac{f^{(2)}}{2} \frac{\bar{v}^2}{2(1 + \bar{n} + f^{(1)}\bar{v})} f(\bar{v}).
 \end{aligned} \tag{15}$$

This diagram exists in both the stochastic LIF and linear-reset models. It arises from the nonlinear intensity function  $f$  and vanishes almost everywhere if  $f$  is threshold linear. The vertex  $\circ$  in it carries the factor of  $f^{(2)}/2$ . This contribution impacts the mapping from membrane potential to firing rate. Similarly to above, the approximate cumulant  $\langle\langle v(t)v(t) \rangle\rangle$  is positive in an active state with nondecreasing  $f$ . So if  $f^{(2)} > 0$ , fluctuations will promote activity and vice versa. The curvature also determines the magnitude of this contribution.

Evaluating the one-loop predictions at an equilibrium of  $\bar{v}$ , the stochastic LIF with a threshold-linear intensity function has the one-loop equilibrium:

$$\begin{aligned}
 \bar{v} &= \frac{1}{10} (1 + \sqrt{1 + 80E}), \\
 \bar{n} &= \lfloor \bar{v} - 1 \rfloor_+.
 \end{aligned} \tag{16}$$

Comparing the one-loop and mean-field predictions, we see that the negative vertex factor leads to a suppression of activity [Fig. 1(c), dotted line vs dashed line; Fig. 1(d), black line]. This occurs because the nonlinear spike reset negatively couples the mean membrane voltage to joint fluctuations in the spikes and membrane voltage.

We can also calculate the rate of the threshold-linear stochastic LIF neuron exactly. The spike train is a renewal process because of the nonlinear reset mechanism. Standard results of renewal theory expose its rate [74]. With a constant drive  $E$ , the membrane voltage evolves after a spike at time  $t$  as  $v(t+s) = E[1 - \exp(-s)]$ , with  $v(t) = 0$ . The time-averaged firing rate is the inverse of the mean interspike interval:  $\langle \dot{n} \rangle = 1/\langle s \rangle$ . For threshold-linear  $f$ , the mean interspike interval is

$$\langle s \rangle = \ln\left(\frac{E}{E-1}\right) + \left(\frac{E-1}{e}\right)^{1-E} \gamma(E-1, E-1). \tag{17}$$

$\gamma(x, y)$  is the lower incomplete  $\gamma$  function. The term  $\ln(E/E - 1)$  is the time for  $v(t)$  to reach the threshold value of 1; the second term is the mean first spike time after that. The one-loop prediction matches the true  $f$ - $I$  curve better than the mean-field theory [Fig. 1(c)].

In the uncoupled linear-reset model, the one-loop equations of motion are

$$\begin{aligned} 0 &= \partial_t \bar{v} + \bar{v} + r\bar{n} - E, \\ 0 &= \bar{n} - f(\bar{v}) - \text{loop correction} \end{aligned} \quad (18)$$

The linear-reset model has the same four types of propagator as the stochastic LIF, although their definitions differ between the two models due to the different spike reset mechanisms (Appendix F). The matched linear-reset model [ $f_M(v) = v[v - 1]_+$ , with  $r = 1$ ] has the one-loop equilibrium:

$$\begin{aligned} \bar{v} &= \frac{1}{8}(-1 + \sqrt{64E + 17}), \\ \bar{n} &= \left(\bar{v} + \frac{1}{4}\right)[\bar{v} - 1]_+. \end{aligned} \quad (19)$$

The term  $\frac{1}{4}[\bar{v} - 1]_+$  is the result of the loop correction to the firing rate, leading to an increase in the rate compared to the mean-field rate. The expected membrane potential still reaches the threshold value  $\bar{v} = 1$  at  $E = 1$ , and the one-loop membrane voltage is greater than the mean-field membrane voltage  $\sqrt{E}$ . Close to threshold ( $E = 1 + \epsilon$ ), the one-loop voltage is  $\bar{v} = 1 + 4\epsilon/9 + \mathcal{O}(\epsilon^2)$ , while for the mean-field theory,  $\bar{v} = 1 + \epsilon/2 + \mathcal{O}(\epsilon^2)$ . Similarly, for large  $E$  the one-loop membrane potential is approximately  $1 + \sqrt{E}$ .

In summary, fluctuations suppress activity in the stochastic LIF neuron because the nonlinear spike reset negatively couples the mean membrane voltage to joint spike-voltage fluctuations. In the linear-reset model, the only nonlinearity arises from the intensity function. To match the mean-field voltage of the stochastic LIF neuron, the linear neuron's intensity function is  $f_M(v) = vf(v)$ , where  $f(v)$  is the stochastic LIF intensity function. Since  $f(v)$  had non-negative curvature,  $f_M$  has positive curvature. So, fluctuations of the membrane voltage promote spiking activity in the matched linear-reset model. A stochastic LIF network with a nonlinear intensity function may have contributions from both diagrams, so that the two nonlinearities compete to determine whether fluctuations suppress or promote activity.

#### IV. HOMOGENEOUS NETWORKS

Biological neural networks are coupled. We will seek a low-dimensional description of the population activity that accounts for synaptic coupling. Here, we study the simplest case: networks where the connectivity between neurons is homogeneous, so we take the synaptic weights between neurons from a distribution with negligible second- and higher-order cumulants. We assume that the mean synaptic weight is  $\mathcal{O}(1/N)$  so the total synaptic weight onto a neuron is  $\mathcal{O}(1)$ . An exemplar of this case is a network with weak ( $J_{ij} \sim 1/N$ ) but potentially dense (connection probability  $\sim 1$ ) connections (Appendix C). To examine the interaction between synaptic connectivity, subthreshold dynamics, and stochastic spike emission in shaping network activity, we will average the partition functional for the activity (equivalently, average the moment generating functional) over realizations of the synaptic connectivity (Appendix C). In the limit of large  $N$ , the density factorizes over the neurons, yielding the partition functional:

$$Z^* = \int \mathcal{D}v \int \mathcal{D}\dot{n} \int \mathcal{D}\bar{v} \int \mathcal{D}\bar{n} \exp\left(-\bar{v}^T(\partial_t v + v - E - J * \langle \dot{n} \rangle + \dot{n}v) - \bar{n}^T \dot{n} + [\exp(\bar{n}) - 1]^T f\right). \quad (20)$$

The result is a population of independent stochastic LIF neurons, each receiving a self-consistent mean-field input  $J * \langle \dot{n} \rangle$ , where  $\langle \dot{n} \rangle$  is the population-averaged spike train. For self-averaging connectivity, the result describes the typical behavior of an individual network and the population average  $\langle \dot{n} \rangle$  matches the ensemble averaged rate. Since the density factorizes, we drop the neuron index. Robert and Touboul proved convergence to these mean-field dynamics [75]. The connectivity has been reduced to its mean  $J$ , which would be equivalent to assuming a network with all-to-all connectivity.  $J$  can be either positive or negative. If the connectivity had non-negligible higher cumulants, these would give rise to corresponding

fluctuations in the membrane potential (Appendix C). This population-averaged mean-field theory is one dimensional not because the neurons are synchronized, but because they spike independently given a self-consistent mean-field input.

If the network is in an asynchronous state so  $\langle \dot{n} \rangle$  is constant in time, after a spike at time  $t$  the membrane voltage obeys

$$v(t + s) = (E + J\langle \dot{n} \rangle)[1 - \exp(-s)] \quad (21)$$

and the spike train is a renewal process. [We write  $J$  for the integral of the coupling kernel  $J(s)$ .] With a

threshold-linear intensity function, the mean interspike interval is

$$\langle s \rangle = \ln\left(\frac{C}{C-1}\right) + \left(\frac{C-1}{e}\right)^{1-C} \gamma(C-1, C-1), \quad (22)$$

where  $C = E + J\langle \dot{n} \rangle$ . In a stationary state, the rate is the inverse of the mean interspike interval:  $\langle \dot{n} \rangle = 1/\langle s \rangle$ , which allows us to find self-consistent solutions of Eq. (22) numerically.

The mean-field (tree-level) equation of motion for the membrane voltage is

$$0 = \partial_t \bar{v} + \bar{v} + \bar{v} f(\bar{v}) - E - J * f(\bar{v}), \quad (23)$$

with  $f(\bar{v})$  the mean-field approximation of  $\dot{n}$ . As in the  $N$ -dimensional mean-field theory of Eq. (5), this neglects all fluctuations, so we expect that it will not be quantitatively correct. Since the spike trains are conditionally Poisson, those fluctuations are driven by the expected intensity. We thus expect that Eq. (23) should be a good approximation when the true firing rate is low. As we will see below, it can provide a good qualitative description of the population dynamics, including bifurcations from quiescence. The leading-order description of fluctuations is given by the one-loop equations of motion:

$$\begin{aligned} 0 &= \partial_t \bar{v} + \bar{v} + \bar{v} \bar{n} - E - J * \bar{n} + \text{loop}, \\ 0 &= \bar{n} - f(\bar{v}) - \text{loop}. \end{aligned} \quad (24)$$

The one-loop contributions are given by Eqs. (14) and (15).

## V. BISTABLE ACTIVITY IN HOMOGENEOUS NETWORKS

With a threshold-linear  $f$ ,  $f(v) = [v - 1]_+$ , and pulse coupling,  $J(s) = J\delta(s)$ , there are three possible steady states of Eq. (23). The first is  $\bar{v} = E$ , which exists if  $E < 1$ . There are two other possible steady states at  $v > 1$ :

$$\begin{aligned} \bar{v}_\pm &= \frac{J \pm \sqrt{J^2 + 4(E - J)}}{2}, \\ \bar{n}_\pm &= [\bar{v}_\pm - 1]_+, \end{aligned} \quad (25)$$

which both exist if

$$E < 1 \quad \text{and} \quad J > 2 + 2\sqrt{1 - E}. \quad (26)$$

Whenever it exists,  $\bar{v}_-$  ( $\bar{v}_+$ ) is unstable (stable). If  $E > 1$ , only  $\bar{v}_+$  exists. With  $J > 2$  and  $[J(4 - J)/4] \leq E < 1$ , both steady states exist and the firing rates are thus bistable, with

$\bar{v}_-$  providing a separatrix between the attractors  $v \rightarrow E$  and  $v \rightarrow \bar{v}_+$ . The mean-field theory has two saddle node bifurcation curves, where the unstable fixed point  $\bar{v}_-$  meets either  $\bar{v} = E$  or  $\bar{v}_+$  [Fig. 2(a)].

These bifurcations also appear in the underlying stochastic spiking model. We simulated a network of 100 stochastic LIF neurons [Eq. (2)] with Erdős-Rényi connectivity ( $p = 0.5$ ) with different values of the baseline drive  $E$  and coupling strength  $J$  [marked in Fig. 2(a)]. At times 5 and 15, we applied pulse perturbations to the baseline drive and observed monostable or bistable behavior matching the predictions of the phase diagram [Figs. 2(b)–2(d)].

The mean-field theory neglects all fluctuations in the spiking activity. Because of the nonlinear spike-voltage coupling imparted by the reset mechanism, those fluctuations can impact the firing rate. To determine the magnitude of fluctuation corrections, we computed bifurcation diagrams of the exact firing rate [Eq. (22); see Figs. 2(e) and 2(f)]. The mean-field theory systematically overestimates the true firing rates. This implies that fluctuations in the activity suppress firing.

Similarly to the uncoupled neuron, the impact of fluctuations can be explicitly described by loop corrections to the mean-field dynamics [Eq. (24)]. To one loop, equilibria of  $\bar{v}$  are  $\bar{v} = E$  (if  $\bar{v} < 1$ ) and

$$\begin{aligned} \bar{v}_\pm &= \frac{1}{10} \left( 1 + 4J \pm \sqrt{1 + 80E + 8J(2J - 9)} \right), \\ \bar{n}_\pm &= [\bar{v}_\pm - 1]_+, \end{aligned} \quad (27)$$

if  $\bar{v} > 1$ . At one loop, both equilibria of  $\bar{v}$  exist if

$$E < 1 \quad \text{and} \quad J > \frac{9}{4} + \sqrt{5(1 - E)}. \quad (28)$$

In the model with linear resets and a threshold-linear intensity function, the mean-field theory is linear in both the subthreshold and suprathreshold regimes and does not exhibit bistability. The classic activity equations can have bistable regimes so long as the nonlinearity saturates; see, e.g., Ref. [8]. Here, bistability is due to the nonlinear coupling between the spiking and membrane voltage.

The stochastic spiking network may not exhibit true bistability in the bistable regime of the deterministic mean-field or one-loop approximations. Rather, the quiescent state should be truly stable, while the active state is metastable. Fluctuations in the spiking activity may drive the network into the quiescent state. In the quiescent state, there are no fluctuations since all  $n$ -point correlation functions are sourced by the intensity  $f(v)$ , which we took to be 0 for  $v < 1$ . If the nonlinearity  $f(v)$  were small but finite for  $v < 0$ , then fluctuations could be maintained in the quiescent state and both would be metastable. The slope

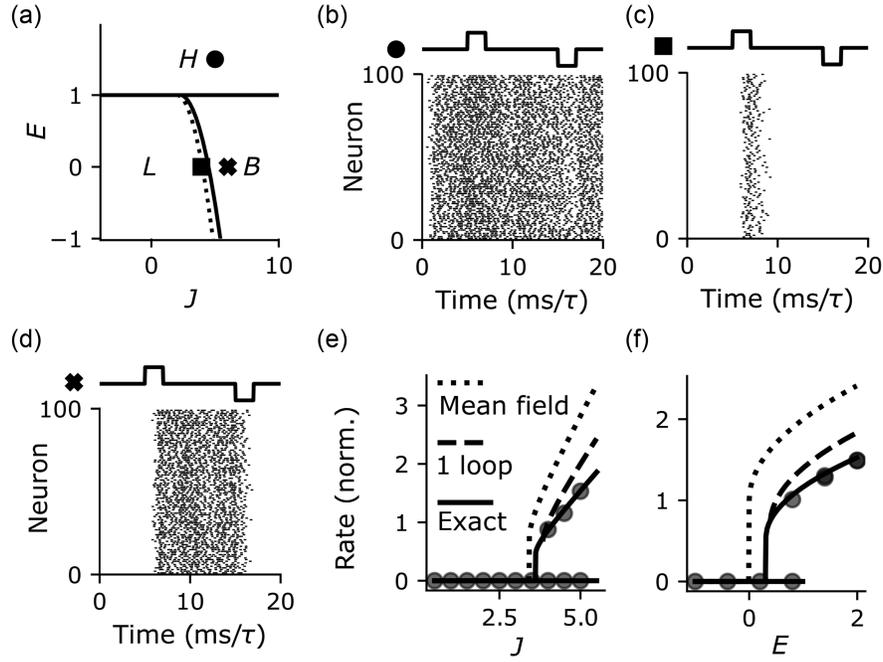


FIG. 2. Bistable activity in homogeneous networks. (a) Phase diagram of the mean-field theory, Eq. (23), in the input ( $E$ ) vs coupling ( $J$ ) plane. There are three possible states: low activity ( $L$ ), high activity ( $H$ ), and bistability ( $B$ ). (b)–(d) Raster plots of a homogeneous network’s activity at the parameter locations marked in (a). At  $t = 5$  and  $t = 15$ , perturbations of amplitude 2 and duration 2 are applied to the drive  $E$  (top). (e) Bifurcation curve in  $J$  with  $E = 1/2$ . (f) Bifurcation plot in  $E$  with  $J = 4$ . Gray circles: simulation. Black dotted line: mean field theory of Eq. (23). Black dashed line: one-loop theory of Eq. (24). Black solid line: the exact rate of the disorder-averaged system, using the numerical self-consistent solution of Eq. (22). The simulated network has Erdős-Rényi connectivity. Simulated network parameters:  $N = 100$ ,  $p = 0.5$ . All nonzero connections have the same weight,  $J/(pN)$ .

of the intensity function at threshold can also play a key role in metastability of the population activity [75].

## VI. MULTIPLE CELL TYPES

Biological neural networks are composed of diverse types of neurons with cell-type-specific connectivity; see, e.g., Ref. [76–83]. Motivated by this, we consider a

network with  $M$  populations, which impose a block structure on the connectivity matrix  $J$ . The average over the connectivity proceeds as for the single population, with an order parameter for each population’s mean activity. This yields an  $M$ -dimensional mean-field theory. In the large- $N$  limit, the partition functional is

$$Z^* = \int \mathcal{D}\mathbf{v} \int \mathcal{D}\dot{\mathbf{n}} \int \mathcal{D}\tilde{\mathbf{v}} \int \mathcal{D}\tilde{\mathbf{n}} \exp \sum_{\alpha=1}^M \left[ -\tilde{v}_\alpha^T \left( \partial_t v_\alpha + v_\alpha - E_\alpha - \sum_{\beta=1}^M J_{\alpha\beta} * \langle \dot{n}_\beta \rangle + \dot{n}_\alpha v_\alpha \right) - \tilde{n}_\alpha^T \dot{n}_\alpha + [\exp(\tilde{n}_\alpha) - 1]^T f(v_\alpha) \right]. \quad (29)$$

For self-averaging networks, the density factorizes over the populations and neurons so the neurons again spike independently given a self-consistent mean-field input. The typical spike train of population  $\alpha$  ( $\alpha \in [M]$ ) is an inhomogeneous Poisson process. If the population-averaged activities  $\langle \dot{n}_\alpha \rangle$  are constant in time, the mean first passage times are

$$\langle s_\alpha \rangle = \ln \left( \frac{C_\alpha}{C_\alpha - 1} \right) + \left( \frac{C_\alpha - 1}{e} \right)^{1-C_\alpha} \gamma(C_\alpha - 1, C_\alpha - 1), \quad (30)$$

where  $C_\alpha = E + \sum_{\beta=1}^M J_{\alpha\beta} \langle \dot{n}_\beta \rangle$ . In a stationary state, the rate is  $\langle \dot{n}_\alpha \rangle = 1/\langle s_\alpha \rangle$ . The mean-field approximation of the membrane voltages is

$$\partial_t v_\alpha = -v_\alpha - v_\alpha f(v_\alpha) + E_\alpha + \sum_{\beta} J_{\alpha\beta} * f(v_\beta). \quad (31)$$

The one-loop equations of motion, similarly, are given by accounting for the input across populations in Eq. (24).

## VII. BISTABLE ACTIVITY IN EXCITATORY-INHIBITORY NETWORKS

Here, we consider the classic excitatory-inhibitory network with pulse coupling and mean connection strengths,

$$\begin{pmatrix} J_{EE} & J_{EI} \\ J_{IE} & J_{II} \end{pmatrix} = \begin{pmatrix} J & -gJ \\ J & -gJ \end{pmatrix}, \quad (32)$$

as in Refs. [20,21] [Fig. 3(a)]. With input  $E$  to both populations, the mean rates of the excitatory and inhibitory populations are equal since they receive the same external and recurrent inputs. The self-consistent fixed points with positive rates are the same as those in the single-population network with the replacement  $J \rightarrow J(1-g)$  [Eq. (25) for the mean-field theory, Eq. (27) to one loop]. In the mean-field theory, both fixed points exist if

$$\begin{aligned} E < 1 \quad \text{and} \quad J > 2 \\ \text{and} \quad g \leq 1 - \frac{2}{J} \left(1 + \sqrt{1-E}\right). \end{aligned} \quad (33)$$

Here we have highlighted the requirement for  $g$  as a function of the other model parameters;  $J > 2$  is a necessary condition for bistability in the mean-field theory, but depending on the values of  $g$  and  $E$  the greatest lower bound for  $J$  may be above 2. With both population voltages under threshold, there is the stable fixed point  $\bar{v} = E$ , if  $E \leq 1$ . If  $E > 1$ , only  $\bar{v}_+$  exists. The Jacobian eigenvalue  $J(1-g) - 2v$  is positive for  $\bar{v}_-$  and negative for  $\bar{v}_+$  root. So if these fixed points exist, the one at higher  $v$  is stable and the other a saddle. Similarly, in the one-loop theory both fixed points exist if

$$\begin{aligned} E < 1 \quad \text{and} \quad J > \frac{9}{4} \\ \text{and} \quad g \leq 1 - \frac{1}{J} \left(\frac{9}{4} - \sqrt{5(1-E)}\right). \end{aligned} \quad (34)$$

As for the single-population network, the existence conditions for these fixed points define saddle node bifurcation curves for the mean-field and one-loop theories [Figs. 3(b)–3(d)]. If the inhibitory coupling strength is sufficiently low, we have the same types of bifurcation curves as in the single-population network [Fig. 3(c)]. If the inhibitory coupling  $g$  is too strong, the only stable equilibrium is the low-rate state [Figs. 3(b) and 3(d)].

These bifurcations also appear in the stochastic spiking network with block-Erdős-Rényi connectivity [Figs. 3(e) and 3(f); network parameters are given in the caption].

## VIII. FLUCTUATIONS

The temporal structure of fluctuations can shape sensory codes [84–86] and determine neural circuit structures

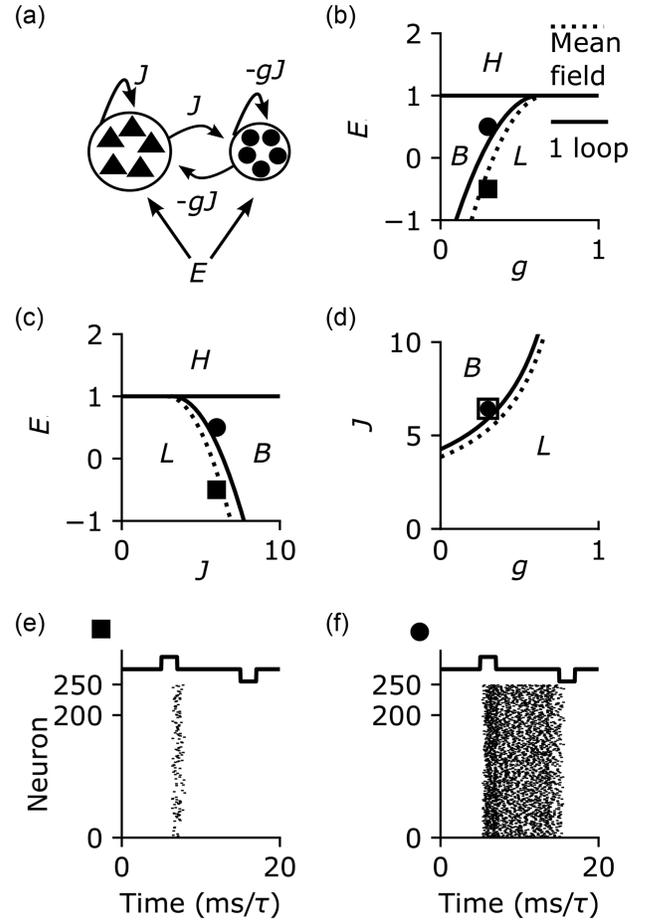


FIG. 3. Bistable activity in excitatory-inhibitory networks. (a) Network diagram. (b) Phase diagram in the input ( $E$ ) vs inhibitory strength ( $g$ ) plane with  $J = 6$ . There are three possible states: low activity ( $L$ ), high activity ( $H$ ), and bistability ( $B$ ). (c) Phase diagram of the two-dimensional mean-field theory, Eq. (31), in the input ( $E$ ) vs coupling strength ( $J$ ) plane with  $g = 0.3$ . (d) Phase diagram in the coupling vs inhibitory strength plane with  $E = 0.5$ . (e),(f) Example simulations with  $(J, g) = (6, 0.3)$ , with  $E = -0.5$  (e) or  $E = 0.5$  (f). The network has a block-Erdős-Rényi structure. Simulated network parameters: population sizes  $(N_e, N_i) = (200, 50)$ , excitatory output connection probabilities  $p_{ee} = p_{ie} = 0.2$ , inhibitory output connection probabilities  $p_{ei} = p_{ii} = 0.8$ . Within each block, all nonzero connections have the same weight, e.g.,  $J/(N_e p_e)$  for nonzero excitatory projections.

through spike timing-dependent plasticity [87–93]. The classic Fokker-Planck mean-field theory of integrate-and-fire networks assumes that the membrane voltages experience a white Gaussian noise [20–22]. The resulting predictions for the spike trains’ power spectra are not white, however, so these predictions are not self-consistent [25]. In the stochastic integrate-and-fire model, the output spike trains also are not white. In the excitatory-inhibitory network, for example, the population-averaged power spectrum exhibits a high-pass shape with a slight resonance [Fig. 4(a), dots]. This is similar to the shape of the power

spectrum of networks of deterministic integrate-and-fire neurons with white noise inputs [23]. The one-loop equations of motion account for Gaussian fluctuations, but do not make any assumptions about their temporal structure. We will next discuss the temporal structure of fluctuations in this Gaussian approximation and the full prediction from renewal theory.

$$p(s_\alpha) = \begin{cases} 0, & s_\alpha \leq \ln \frac{C_\alpha}{C_\alpha - 1} \\ \left[ C_\alpha (1 - e^{-s_\alpha}) - 1 \right] \times \exp \left[ -C_\alpha e^{-s_\alpha} + (C_\alpha - 1) \left( s_\alpha - 1 - \ln \frac{C_\alpha}{C_\alpha - 1} \right) \right], & s_\alpha > \ln \frac{C_\alpha}{C_\alpha - 1}, \end{cases} \quad (35)$$

where  $C_\alpha = E + \sum_\beta J_{\alpha\beta} \langle \dot{n}_\beta \rangle$  and  $\gamma$  is again the lower incomplete  $\gamma$  function. This provides an exact prediction for the interspike interval density in the limit  $N \rightarrow \infty$ , accurate for populations of a few hundred neurons [Fig. 4(b)]. The interspike interval distribution defines the spike train power spectrum  $C(\omega)$  of a renewal process [94]:

$$C(\omega) = \langle \dot{n} \rangle \frac{1 - |p(\omega)|^2}{|1 - p(\omega)|^2}. \quad (36)$$

Together, Eqs. (35) and (36) provide an exact prediction for the typical power spectrum in a large homogeneous network. Computing the Fourier transform  $p(\omega)$  numerically, we see that these predictions are quantitatively accurate in simulations of a few hundred neurons [Fig. 4(a), gray dots vs solid black curve].

For analytic approximations of the power spectrum, we turn to the field-theoretic formulation. If the fluctuations or the nonlinearity is weak, we can expand the density perturbatively around a solution of the deterministic mean-field theory (Appendix D). The connected two-point function of the spike trains can then be calculated diagrammatically. With a threshold-linear intensity function,

$$\langle \langle \dot{n}^2 \rangle \rangle = \text{tree} + \text{loop} + \text{blob} + \mathcal{O}(2 \text{ loops}). \quad (37)$$


The expansion may contain terms with up to infinitely many loops, inducing dependence on  $n$ -point correlation functions of all orders. With an intensity function nonlinear at the mean voltage there would be additional diagrams, containing internal vertices with multiple incoming wavy edges. The same is true for any cumulant of the activity. The simplest approximation of the two-point correlation is the tree-level approximation given by the first diagram of Eq. (37):

The exact mean-field theory, Eq. (29), is of an inhomogeneous Poisson process receiving the self-consistent mean input  $\sum_\beta J_{\alpha\beta} \langle \dot{n}_\beta \rangle$ . Substituting this into Eq. (21) yields the postspike membrane voltage, which defines the intensities  $f(v_\alpha(t))$ . The interspike interval density is  $p(s_\alpha) = f(v_\alpha(s)) \exp[-\int_0^s dt f(v_\alpha(t))]$ . For a threshold-linear intensity function, it is

$$\langle \langle \dot{n}_\alpha^2 \rangle \rangle_0(\omega) = \frac{\bar{v}_\alpha^2 + \omega^2}{4\bar{v}_\alpha^2 + \omega^2} f(\bar{v}_\alpha), \quad (38)$$

where  $\bar{v}_\alpha$  is a solution to the mean-field equation for population  $\alpha$ . At  $\omega = 0$ , this yields  $\langle \langle \dot{n}_\alpha^2 \rangle \rangle_0(0) \approx f(\bar{v}_\alpha)/4$ . For  $\omega \rightarrow \infty$ ,  $\langle \langle \dot{n}^2 \rangle \rangle(\omega) \rightarrow f(\bar{v}_\alpha)$ , the mean-field approximation to the intensity. This simple approximation captures the high-pass nature of the power spectrum but is not quantitatively accurate [Fig. 4(a), dotted line].

The one-loop predictions for the mean membrane voltage and rate account for second-order fluctuations to tree level. For the spike train power spectrum this again corresponds to Eq. (38), but with  $\bar{v}$  a solution to the one-loop equations of motion. This provides a more accurate prediction of the power spectrum [Fig. 4(a), dashed line] due to the improved estimate of the intensity  $f(\bar{v}_\alpha)$ .

As the coupling or input strength brings the network to a bifurcation, the spike train variance  $\langle \langle \dot{n}_\alpha^2 \rangle \rangle_0(0)$  undergoes a sharp transition from 0 in the quiescent state to positive values in the active state [Figs. 4(c)–4(f)]. The transition in the spike train variance follows that in the rate, since all correlation functions are sourced by the intensity  $f(\bar{v})$ .

## IX. INHIBITORY STABILIZATION

In recent years, a body of work has emerged suggesting that mammalian cortices reside in an inhibition-stabilized regime [8,59,61–64]. There are two requirements for an excitatory-inhibitory network to be inhibition stabilized: The network must occupy a stable attractor, but the excitatory population would be unstable on its own. These are difficult to directly test experimentally. Fortunately, inhibition-stabilized fixed points have another signature: paradoxical responses to inhibitory neuron stimulation. In an inhibition-stabilized network, stimulation of the inhibitory neurons leads to a paradoxical reduction of their firing rates [60]. If there are multiple inhibitory subtypes, the net inhibitory input to excitatory neurons decreases upon inhibitory neuron stimulation [95]. The widespread experimental observation of paradoxical

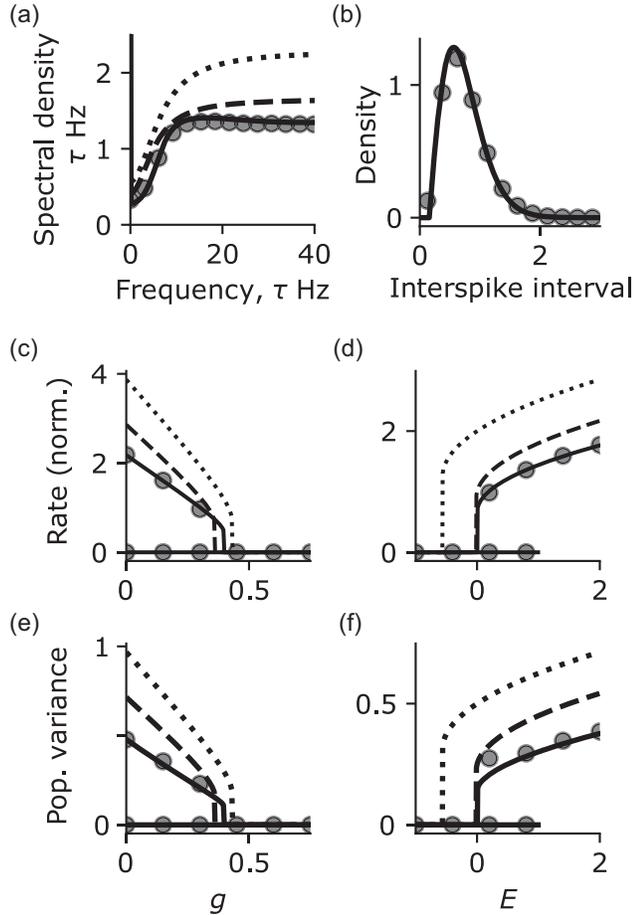


FIG. 4. Fluctuations in an excitatory-inhibitory network with symmetric external inputs,  $E_e = E_i = E$ . (a) Spike train power spectrum with  $(J, g, E) = (6, 0.3, 1.2)$ . Dots: simulation of a network with 200 excitatory and 50 inhibitory neurons (population-averaged power spectrum). Dotted line: the perturbative tree-level approximation (expanded around the deterministic mean-field theory). Dashed line: the tree-level approximation around the one-loop rates. Solid line: the renewal prediction of Eq. (36). (b) Interspike interval density with  $(J, g, E) = (6, 0.3, 1.2)$ . Dots: simulation. Solid line: the renewal prediction of Eq. (35). (c),(d) Bifurcation diagrams for firing rate as a function of  $g$  and as a function of  $E$ . (e),(f) Bifurcation diagrams for spike train variance as a function of  $g$  and as a function of  $E$ . In (c) and (e),  $(J, E) = (6, 0.5)$ . In (d) and (f),  $(J, g) = (6, 0.25)$ . Simulated network parameters: population sizes  $(N_e, N_i) = (200, 50)$ , excitatory output connection probabilities  $p_{ee} = p_{ie} = 0.5$ , inhibitory output connection probabilities  $p_{ei} = p_{ii} = 0.8$ .

responses, and other response patterns consistent with inhibition-stabilized networks, raises the following question: Is inhibitory stabilization a generic property, or does it require fine-tuned parameters toward which cortical networks develop?

The inhibition-stabilized regime, and paradoxical responses as its signature, are predictions of the classic activity equations, Eq. (1). Does an inhibition-stabilized regime exist in the mean-field theory of Eqs. (31) and (32)?

The stability requirements are determined from the Jacobian matrix:

$$\begin{pmatrix} -1 - \bar{f}_e + (J - \bar{v}_e)f_e^{(1)} & -gJf_i^{(1)} \\ Jf_e^{(1)} & -1 - f_i - (gJ + \bar{v}_i)f_i^{(1)} \end{pmatrix}, \quad (39)$$

where  $\bar{f}_\alpha = f(\bar{v}_\alpha)$  and  $f_\alpha^{(1)} = (d/dv)f(v)|_{\bar{v}_\alpha}$ . For a fixed point to be inhibition stabilized, the first element of its Jacobian must be positive (the excitatory-only subnetwork would be unstable), but the maximum real part of its eigenvalues negative (the full network is stable). For the threshold-linear intensity function,  $f^{(1)}(\bar{v}_\alpha) = \theta(\bar{v}_\alpha - 1)$ , where  $\theta(x)$  is the Heaviside step function. This leads to the requirement that for the excitatory subnetwork to be linearly unstable with a positive firing rate,  $1 < \bar{v}_E < J/2$ .

Do paradoxical responses to inhibitory stimulation occur in the stochastic LIF network? To investigate this, we return to the tractable threshold-linear intensity function. We allow the external input to vary between the two populations,  $E = (E, hE)$  [Fig. 5(a)].  $h$  controls the relative strength of the input to the inhibitory population. When both population voltages are above threshold, the mean-field inhibitory and excitatory nullclines are at

$$\begin{aligned} \bar{v}_i^* &= \frac{-gJ + \sqrt{g^2J^2 + 4[J(\bar{v}_e - 1 + g) + hE]}}{2}, \\ \bar{v}_e &= [-(\bar{v}_e^*)^2 + J\bar{v}_e^* - J(1 - g) + E]/gJ. \end{aligned} \quad (40)$$

$\bar{v}_\alpha^*$  is the nullcline of population  $\alpha \in \{e, i\}$ . The supra-threshold inflection point of the excitatory nullcline is at  $\bar{v}_e = J/2$ . An inhibition-stabilized fixed point must thus be on the increasing branch of the excitatory nullcline.  $h$  does not affect the excitatory nullcline but shifts the inhibitory nullcline. An increase in  $h$  will lead to a paradoxical reduction in firing rates if it shifts a stable fixed point to lower  $\bar{v}_i$ . For example, consider the case when there is a single fixed point on the increasing side of the excitatory nullcline, to the left of its peak [Fig. 5(b)]. Increasing  $h$  shifts the inhibitory nullcline up and to the left, moving that fixed point to a lower  $(\bar{v}_e, \bar{v}_i)$ . Depending on the magnitude of the shift, it may also take the dynamics through a bifurcation into a bistable regime. A sufficiently large increase in  $h$  can shift the network into a regime with no excitatory activity, which can also lead to a net decrease in inhibitory rates [Figs. 5(b) and 5(c)].

In what regions of parameter space does an inhibition-stabilized fixed point exist? As discussed above, for the excitatory subnetwork to be unstable, with nonzero excitatory rate, requires that the fixed point be on the middle branch of the excitatory nullcline:  $1 < \bar{v}_E < J/2$ . The inhibitory nullcline is an increasing function of  $\bar{v}_e$ . The excitatory nullcline increases for  $\bar{v}_e^*$  close to 1 and

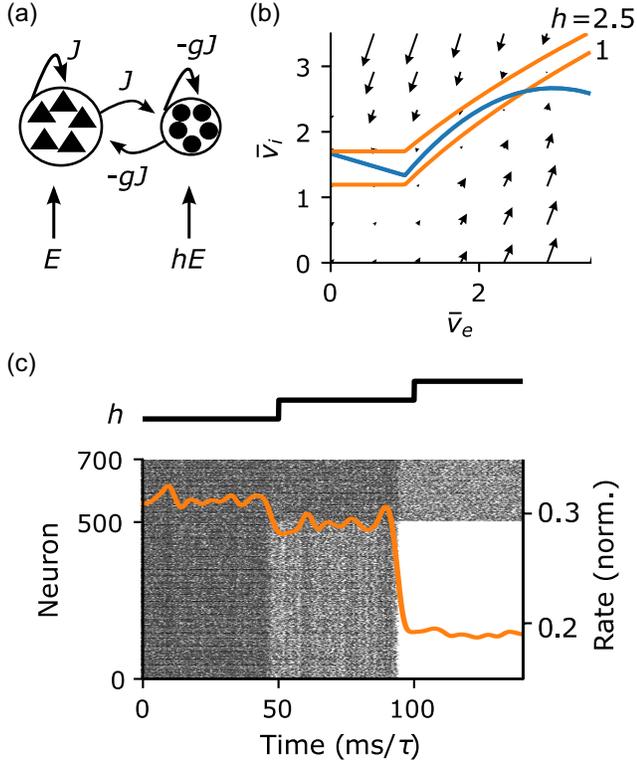


FIG. 5. Paradoxical responses to inhibitory stimulation. (a) Excitatory-inhibitory network with asymmetric drive. (b) Phase diagram and nullclines of the excitatory (blue line) and inhibitory (orange lines) firing rates for the excitatory-inhibitory network with threshold-linear rate functions. (c) Simulation of a block-Erdős-Rényi network with  $p_{ee} = p_{ie} = 0.5$ ,  $p_{ei} = p_{ii} = 0.8$ . At time 0,  $(E, h) = (2, 1)$ . At times 50 and 100,  $h$  increases by  $3/4$ . Orange line: inhibitory population-averaged spike train, smoothed with a Gaussian kernel of width 2 for visualization. Parameters for (b) and (c):  $(J, g, E) = (6, 1/2, 2)$ . Simulated block Erdős-Rényi network parameters as in Fig. 3.

decreases for sufficiently large  $\bar{v}_e^*$ . At threshold ( $\bar{v}_e = 1$ ) the inhibitory nullcline must be below the excitatory nullcline:

$$\sqrt{(gJ)^2 + 4gJ + 4hE} < \frac{2}{gJ}(E + gJ - 1) + gJ. \quad (41)$$

If  $h = 1$ , this requirement imposes that  $E > 1$ ; at  $E = 1$  the two sides are equal, and the difference of the two sides grows as  $\sqrt{E}$ .

The peak of the excitatory nullcline is at  $\bar{v}_e^* = J/2$ . At the peak of the excitatory nullcline,  $\bar{v}_i = [J^2/4 - J(1-g) + E]/gJ$ . At  $\bar{v}_e = J/2$ , the inhibitory nullcline should be above the excitatory nullcline:

$$\begin{aligned} & \sqrt{(gJ)^2 + 4gJ + 2J(J-2) + 4hE} \\ & > \frac{2}{gJ} \left( \frac{J^2}{4} - J(1-g) + E \right) + gJ. \end{aligned} \quad (42)$$

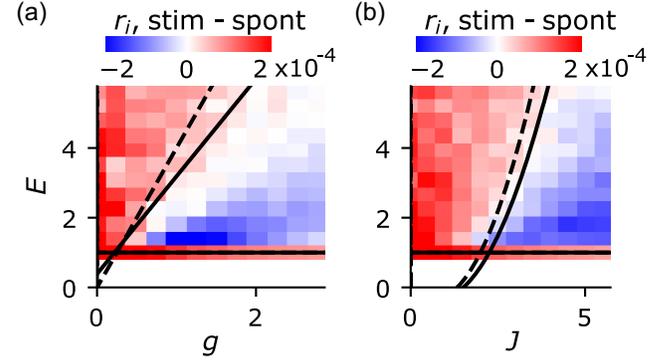


FIG. 6. Phase diagrams for paradoxical responses to inhibitory stimulation. (a) Boundaries of the paradoxical response region with  $J = 4$ . Dashed line: mean-field theory, Eq. (42). Solid line: one-loop theory, Eq. (45). Color: simulation. Each simulation lasts for 200 time units; at time 100, the inhibitory drive switches from  $hE = E$  to  $hE = E + 0.1$ . (b) As in (a), with  $g = 2$ . Simulated block-Erdős-Rényi network parameters as in Fig. 3.

Together, Eqs. (41) and (42) provide sufficient conditions for a paradoxical response to inhibitory stimulation in the mean-field theory. At fixed drive  $E$ , they predict a paradoxical response for sufficiently large  $J$  or  $g$ . For stronger  $E$ , these minimal couplings increase [Figs. 6(a) and 6(b), dashed line].

To estimate how fluctuations impact inhibitory stabilization, we compute the one-loop nullclines (each with  $v_i$  as a function of  $v_e$ ):

$$\begin{aligned} \bar{v}_i^* &= \frac{1 - 4gJ + \sqrt{1 + 80hE + 8gJ(9 + 2gJ) + 80J(v_e - 1)}}{10}, \\ v_i &= [4(E - J + gJ) + \bar{v}_e^*(4J + 1) - 5(\bar{v}_e^*)^2]/4gJ. \end{aligned} \quad (43)$$

The inflection point of the one-loop  $\bar{v}_e$  nullcline is at  $\bar{v}_e^* = (1 + 4J)/10$ . At one loop, for the inhibitory nullcline to be below the excitatory nullcline at threshold requires

$$10E > 10 + gJ \left[ -9 - 4gJ + \sqrt{1 + 80Eh + 8gJ(9 + 2gJ)} \right]; \quad (44)$$

for the inhibitory nullcline to be above the excitatory nullcline at  $\bar{v}_e = (1 + 4J)/10$  requires

$$\begin{aligned} 1 + 80E &< 8J \left( 9 - 2J - g \left\{ 9 + 4gJ \right. \right. \\ & \left. \left. - \sqrt{1 + 80Eh + 8J[-9 + 4J + g(9 + 2gJ)]} \right\} \right). \end{aligned} \quad (45)$$

For fixed  $E$  and  $J$ , this one-loop boundary requires a higher  $g$  (stronger inhibition) than the mean-field boundary, better matching the transition observed in simulations [Fig. 6(a), dashed line vs solid line]. Similarly, for fixed  $E$  and  $g$ , the

one-loop boundary is at higher  $J$  (stronger coupling) than the mean-field boundary [Fig. 6(b), dashed line vs solid line]. Together, this comparison indicates that fluctuations shift the region of paradoxical responses to more strongly coupled networks. This comports with the role of fluctuations in suppressing activity.

A paradoxical response could also occur from other dynamical regimes than the single fixed point on the decreasing branch of the excitatory nullcline, such as from a bistable regime. To test whether the underlying spiking model exhibits paradoxical responses, we simulated excitatory-inhibitory networks while varying  $J$  and  $g$ . For each network, we applied a perturbation of amplitude 0.1 to the inhibitory population's input and computed the inhibitory population's average firing rate before and after the perturbation. With  $J$  fixed and varying  $g$ , we observed paradoxical responses for sufficiently large  $g$  [Fig. 6(a), blue]. Similarly, with  $g$  fixed and varying  $J$ , we observed paradoxical responses for sufficiently large  $J$  [Fig. 6(b), blue]. The one-loop predictions better match the region of paradoxical responses than the mean-field predictions (Fig. 6).

## X. DISCUSSION

We constructed a path integral representation for the joint probability density functional of the membrane voltage and spike trains of a network of stochastic LIF neurons, Eq. (4). This exposed a simple deterministic mean-field theory for spiking networks: activity equations with an additional rate-dependent leak arising from the spike resetting [Eq. (5)]. It also exposed fluctuation corrections to the mean-field theory, arising from the two nonlinearities of the intensity function and spike reset [Eq. (24)], the latter of which suppresses activity (Fig. 1). These  $N$ -dimensional systems expose predictions for the activity that depend on a particular connectivity  $\mathbf{J}$ . Large-scale electron microscopy is now revealing such wiring diagrams see, e.g., Refs. [96–108]. Predicting the microscopic dynamics of even deterministic, threshold-linear neuronal networks is challenging [109–112]. Statistical approaches focusing on stochastic models allow the prediction of correlations between specific neurons' activity through linear response or Hawkes theory and its generalizations [55,113–118]. Equation (4) provides a starting point for making such predictions in a stochastic LIF network.

Here, we instead used the path integral representation to derive a population-averaged stochastic field theory for large networks with homogeneous coupling, including multipopulation systems like excitatory-inhibitory networks. That stochastic field theory was of the form of a renewal process with a self-consistent input [Eqs. (20) and (29)]. Robert and Touboul studied the homogeneous stochastic LIF network rigorously [75]. They proved that the mean-field process, Eq. (20), can have one or several invariant densities depending on the form of the firing function. The stochastic field theory admits low-dimensional

mean-field and loop approximations of the voltage and rate as simple functions of the model's parameters. Using these approximations, we demonstrated bistability of the deterministic mean-field theory and its extension to the stochastic system, and studied the contributions of recurrent inhibition and spike resetting to stabilizing network activity. We also found that fluctuations suppress activity through the spike reset also in coupled networks [Figs. 2(e), 2(f), and 4(c)–4(e)]. Excitatory-inhibitory networks of deterministic integrate-and-fire neurons can also exhibit bistable equilibrium rates if the inhibition is not too strong [21,119]. The field-theoretic description here does not rely on a white noise approximation for the membrane voltages, but exposes a systematic method for calculating their statistics. It requires here, however, a model with stochastic spike emission.

Deterministic integrate-and-fire networks can also exhibit spatial, temporal, and spatiotemporal transitions [21,120,121]. Temporal, spatial, and spatiotemporal bifurcations are often understood through the classic activity equations [3]. The field theory developed here provides a route to uncovering bifurcations in networks of stochastic integrate-and-fire neurons with more temporal or spatial structure in their interactions, as well as investigating the impact of spiking fluctuations on such transitions.

In the classic activity equations [e.g., Eq. (1)], recurrent inhibition is necessary to stabilize strongly coupled networks [8,59]. An inhibition-stabilized regime can be exposed by a paradoxical reduction of inhibitory activity after inhibitory stimulation [60]. We calculated the phase diagram for paradoxical responses in stochastic LIF networks, and found that an inhibition-stabilized regime exists in wide regions of parameter space (Figs. 5 and 6). This suggests a generic mechanism underlying the observation of paradoxical responses widely in mammalian cortex [61–64].

There are two complementary approaches to our focus on the density functional of sample paths  $p[\mathbf{v}(t), \dot{\mathbf{n}}(t)]$  for the stochastic LIF model. These complementary approaches focus on the time-dependent probability density function of the membrane voltages  $p(v, t)$  across a population of neurons [18]. In the  $N \rightarrow \infty$  limit and with  $J_{ij} \sim 1/N$ , the population density of membrane voltages in a stochastic LIF network obeys a Volterra integral equation [29,30]. That integral equation can also be written as a partial differential equation, which rigorously exposes the stochastic stability of the population densities in a mean-field limit [122–125]. A finite-size analysis introduces a stochastic term to the population density equations [34,126]. Alternatively, moments for finite-size networks can be analyzed through a replica mean-field approach [127–129]. The path integral approach also exposes a finite-size mean-field theory [Eq. (5)]. Fluctuation corrections to that finite- $N$  mean-field theory can be obtained in the same way as for the large- $N$ , connectivity-averaged system.

The field-theoretic approach is practical and flexible. It exposes simple analytic approximations for any cumulant

of the membrane voltages and/or spike trains via diagrammatic methods, is amenable to finite-size corrections, and applies readily to other models such as those with temporal synaptic interactions, spatially dependent connectivity, conductance-based or strong  $\mathcal{O}(1/\sqrt{N})$  synapses, and additional nonlinearities in the single-neuron dynamics.

A previous version of this article, now retracted [130], contained errors described in the retraction notice [131]. These errors have been corrected in the current article.

Related code can be found at Ref. [132].

### ACKNOWLEDGMENTS

I thank Michael A. Buice, Brent Doiron, and Stefano Recanatani for helpful feedback.

### APPENDIX A: JOINT PROBABILITY DENSITY FUNCTIONAL

We will construct the joint probability density of the membrane voltages and spike trains using the response variable path integral formalism [37–40], reviewed in Refs. [133–135]. We will use boldface lowercase variables for vectors and boldface capital letters for matrices and operators. Given the membrane voltages  $v_{it}$ , we will require that the spikes generated in the network are conditionally independent across neurons  $i$  and time points  $t$ . Here we take the model to be already nondimensionalized, so that time is measured in units of the membrane time constant  $\tau$  and the voltage resets to 0 after a spike. The joint probability density of the membrane voltages  $\mathbf{v}$  and the spikes  $d\mathbf{n}$ , conditioned on the stochastic spike generation, is

$$p(\mathbf{v}, \mathbf{n} | \eta) = \prod_{i=1}^N \prod_{t=1}^{T-1} \delta \left( \frac{dv_{it}}{dt} + v_{it} + \frac{dn_{i,t-1}}{dt} v_{it} - E_{it} - \sum_{j,s} J_{ijs} dn_{j,t-s} \right) \delta(dn_{it} - \eta_{it}). \quad (\text{A1})$$

Here,  $\eta_{it} \sim \text{Bernoulli}[f(v_{it})dt]$ . Introducing the Fourier representation of the  $\delta$  functions and marginalizing over  $\eta$  yields the joint density:

$$p(\mathbf{v}, \mathbf{n}) = \int D\tilde{\mathbf{v}} \int D\tilde{\mathbf{n}} \exp \left[ \sum_{i,t} -\tilde{v}_{it} \left( \frac{dv_{it}}{dt} + v_{it} + \frac{dn_{i,t-1}}{dt} v_{it} - E_{it} - \sum_{j,s} J_{ijs} dn_{j,t-s} \right) - \tilde{n}_{it} dn_{it} + \ln \{ 1 + f(v_{it})dt [\exp(\tilde{n}_{it}) - 1] \} \right]. \quad (\text{A2})$$

The measures are  $D\tilde{\mathbf{n}} = \prod_{i,t} (d\tilde{n}_{it}/2\pi i)$  and  $D\tilde{\mathbf{v}} = \prod_{i,t} (d\tilde{v}_{it}/2\pi i)$ . The integrals over the response variables,  $\tilde{\mathbf{n}}$  and  $\tilde{\mathbf{v}}$ , are along the imaginary axis. The logarithmic term in the exponent is the cumulant-generating function of the Bernoulli spikes. Galves and Löcherbach proved the existence and uniqueness of stationary densities for the discrete-time model with strictly positive intensity function [136].

We next take a continuous-time limit,  $dt \rightarrow 0, T \rightarrow \infty$ , with their product fixed. This defines the functional integration measures  $\mathcal{D}$ . With  $dt \ll 1$ , we expand the natural logarithm in its Taylor series around 1:  $\ln\{1 +$

$f(v_{it})dt[\exp(\tilde{n}_{it}) - 1]\} = f(v_{it})dt[\exp(\tilde{n}_{it}) - 1] + \mathcal{O}[(dt)^2]$ . This yields Eq. (4), with the infinitesimal shift in the reset term in Eq. (2).

### APPENDIX B: ABSOLUTE REFRACTORY PERIOD

With an absolute refractory period of  $T_r$  time steps in the discrete-time dynamics, during which the membrane voltage is clamped within  $\mathcal{O}(dt)$  of 0, the joint density of the spike trains and membrane voltages instead obeys

$$p(\mathbf{v}, \mathbf{n}) = \int D\tilde{\mathbf{v}} \int D\tilde{\mathbf{n}} \exp \left[ \sum_{i,t} -\tilde{v}_{it} \left( \frac{dv_{it}}{dt} + v_{it} + \sum_{s=1}^{T_r} \frac{dn_{i,t-s}}{dt} v_{it} - E_{it} - \sum_{j,s} J_{ijs} dn_{j,t-s} \right) - \tilde{n}_{it} dn_{it} + \ln \{ 1 + f(v_{it})dt [\exp(\tilde{n}_{it}) - 1] \} \right]. \quad (\text{B1})$$

This presents some complication in the continuous-time limit: the refractory term diverges when written as a convolution.

One alternative would be to incorporate a strong, negative self-coupling in diagonal elements of  $\mathbf{J}$ . While not strictly an absolute refractory period, this may mimic its effects. This would affect the definition of the mean-field theory and propagators, but would not give rise to new types of fluctuation correction (no new vertices; see Appendix D 1).

Another alternative is to consider an absolute refractory period in which the membrane voltage is not clamped at the reset voltage. Rather, we can require that

- (1) during the absolute refractory period, the spike probability is 0 and
- (2) at the end of the absolute refractory period, the membrane voltage is reset to the reset voltage.

This yields the discrete-time density:

$$p(\mathbf{v}, \mathbf{n}) = \int D\tilde{\mathbf{v}} \int D\tilde{\mathbf{n}} \exp \left\{ \sum_{i,t} -\tilde{v}_{it} \left( \frac{dv_{it}}{dt} + v_{it} + \frac{dn_{i,t-T_r}}{dt} v_{it} - E_{it} - \sum_{j,s} J_{ijs} dn_{j,t-s} \right) - \tilde{n}_{it} dn_{it} + \ln \left[ 1 + f(v_{it}) dt [\exp(\tilde{n}_{it}) - 1] \left( 1 - \sum_{s=1}^{T_r} dn_{i,t-s} \right) \right] \right\}, \quad (\text{B2})$$

from which a continuum limit can be taken straightforwardly, yielding a spike reset term  $\dot{n}_i(t - \tau_r)v_i(t)$  and spike intensity  $f(v_i)(1 - \dot{n}_i * B)$ , with the rectangular function  $B(t) = \theta(t) - \theta(t - \tau_r)$  and refractory period  $\tau_r = T_r dt$ . This introduces a new state dependence to the intensity, which would give rise to new types of fluctuation correction.

### APPENDIX C: CONNECTIVITY-AVERAGED DENSITY

To examine the interaction between synaptic connectivity, subthreshold dynamics, and stochastic spike emission in shaping network activity, we will average the partition functional for the activity (equivalently, average the moment generating functional) over the synaptic connectivity. This is a standard exercise in statistical field theory [135], relying on the assumption that the system is self-averaging with respect to the connectivity: that is, that the average over realizations of  $\mathbf{J}$  will give us an accurate description of a single large system. The connectivity-averaged partition functional is

$$Z = \int D\mathbf{v} \int D\dot{\mathbf{n}} \int D\tilde{\mathbf{v}} \int D\tilde{\mathbf{n}} \int D\mathbf{J} p[\mathbf{v}, \dot{\mathbf{n}}, \tilde{\mathbf{v}}, \tilde{\mathbf{n}} | \mathbf{J}] p(\mathbf{J}). \quad (\text{C1})$$

Since the action  $S$  is linear in  $\mathbf{J}$ , a cumulant-generating function for  $\mathbf{J}$  appears in  $Z$ :

$$\begin{aligned} Z &= \int D\mathbf{v} \int D\dot{\mathbf{n}} \int D\tilde{\mathbf{v}} \int D\tilde{\mathbf{n}} \exp - \left( \tilde{\mathbf{v}}^T (\partial_t \mathbf{v} + \mathbf{v} - \mathbf{E} + \dot{\mathbf{n}}\mathbf{v}) + \tilde{\mathbf{n}}^T \dot{\mathbf{n}} - [\exp(\tilde{\mathbf{n}}) - 1]^T \mathbf{f} \right) \int D\mathbf{J} p(\mathbf{J}) \exp(\tilde{\mathbf{v}}^T \mathbf{J} * \dot{\mathbf{n}}) \\ &= \int D\mathbf{v} \int D\dot{\mathbf{n}} \int D\tilde{\mathbf{v}} \int D\tilde{\mathbf{n}} \exp - \left( \tilde{\mathbf{v}}^T (\partial_t \mathbf{v} + \mathbf{v} - \mathbf{E} + \dot{\mathbf{n}}\mathbf{v}) - W_{\mathbf{J}}(\tilde{\mathbf{v}}^T \dot{\mathbf{n}}) + \tilde{\mathbf{n}}^T \dot{\mathbf{n}} - [\exp(\tilde{\mathbf{n}}) - 1]^T \mathbf{f} \right), \end{aligned} \quad (\text{C2})$$

where

$$W_{\mathbf{J}}(\tilde{\mathbf{v}}^T \dot{\mathbf{n}}) = \ln \int D\mathbf{J} p(\mathbf{J}) \exp \left( \int dt \int ds \sum_{i,j} \tilde{v}_i(t) J_{ij}(s) \dot{n}_j(t-s) \right). \quad (\text{C3})$$

Because of this, each cumulant of  $\mathbf{J}$  gives rise to a corresponding cumulant in the connectivity-averaged partition functional for the activity. We have overloaded notation here, writing  $p(\mathbf{J})$  for the distribution of the synaptic weight matrix while also letting  $\mathbf{J}$  be a function of the time lag. This notation assumes that  $J_{ij}(s) = J_{ij} G_{ij}(s)$  for some matrix of unit-norm kernels  $\mathbf{G}$ , which we leave implicit. The connectivity gives rise to an effective noise in the membrane voltage. Each cumulant of the connectivity gives rise to a cumulant of the same order in the effective noise.

For example, consider an Erdős-Rényi network with connection probability  $p$  and synaptic weight  $J$  for the nonzero connections, with  $J(s) = J\delta(s)$ . The distribution  $p(\mathbf{J})$  factorizes over the weights; the cumulant-generating functional for an individual synaptic weight is

$$W_J(x) = \ln \{ 1 + p[-1 + \exp(Jx)] \} \quad (\text{C4})$$

and cumulants of the synaptic weights  $J_{ij}$  obey the recursion relation:

$$\langle J_{ij} \rangle = Jp, \quad \langle\langle J_{ij}^n \rangle\rangle = Jp(1-p) \frac{d}{dp} \langle\langle J_{ij}^{n-1} \rangle\rangle, \quad n \geq 2. \quad (\text{C5})$$

If the connection probability  $p$  and weight  $J$  are both of order 1, the synaptic weights will have non-negligible cumulants of all orders. If the synaptic weights are of order 1 and the connectivity sparse,  $p \sim 1/N$ , the cumulants

are  $\langle\langle J_{ij}^n \rangle\rangle = J^n p + \mathcal{O}(1/N^2)$ . If  $J > 1$ , higher cumulants of the connectivity will dominate, giving rise to higher-order cumulants in the effective noise of the membrane voltage. In contrast, if  $J \sim 1/N$  and  $p \sim 1$ , then  $\langle\langle J_{ij}^n \rangle\rangle \sim N^{-n}$ , so in a large network, the first cumulant of the connectivity dominates.

Here, we consider that simple case where  $\mathbf{J}$  has only a first cumulant. Let  $\langle J_{ij}(s) \rangle_{\mathbf{J}} = J(s)/N$ . The average over the connectivity yields

$$Z = \int \mathcal{D}\mathbf{v} \int \mathcal{D}\mathbf{\dot{n}} \int \mathcal{D}\tilde{\mathbf{v}} \int \mathcal{D}\tilde{\mathbf{n}} \exp \sum_i \left[ -\tilde{v}_i^T \left( \partial_t v_i + v_i + \dot{n}_i v_i - E_i - \frac{1}{N} \mathbf{J} * \sum_j \dot{n}_j \right) - \tilde{n}_i^T \dot{n}_i + (e^{\tilde{n}_i} - 1)^T f(v_i) \right]. \quad (\text{C6})$$

(Here,  $*$  represents scalar temporal convolution.) We would like to examine this partition functional in the limit of a large network. Let  $R = (1/N) \sum_j \mathbf{J} * \dot{n}_j$ ; we will enforce this by integrating against  $\delta(NR - \mathbf{J} * \sum_j \dot{n}_j)$ . With the Fourier representation of that  $\delta$  function, we have a generating functional for the auxiliary fields  $R, \tilde{R}$ :

$$Z[k, \tilde{k}] = \int \mathcal{D}R \int \mathcal{D}\tilde{R} \exp \left( -N\tilde{R}^T R + \sum_i \ln Z_i[R, \tilde{R}] + \tilde{k}^T R + k^T \tilde{R} \right),$$

$$Z_i[R, \tilde{R}] = \int \mathcal{D}v_i \int \mathcal{D}\dot{n}_i \int \mathcal{D}\tilde{v}_i \int \mathcal{D}\tilde{n}_i \exp \left( -\tilde{v}_i^T (\partial_t v_i + v_i - E - R + \dot{n}_i v_i) + \tilde{R}^T (\mathbf{J} * \dot{n}_i) - \tilde{n}_i^T \dot{n}_i + (e^{\tilde{n}_i} - 1)^T f(v_i) \right). \quad (\text{C7})$$

Note that the generating function for the neural dynamics factorizes over the neurons;  $Z_i[R, \tilde{R}]$  does not contain any other indices. So, we will drop the neuron indices and write  $N \ln Z[R, \tilde{R}]$  instead of  $\sum_i \ln Z_i[R, \tilde{R}]$ . For large  $N$ , we evaluate the integrals over the auxiliary fields  $R, \tilde{R}$  by a saddle point approximation. The saddle point equations are

$$0 = -NR^* + N \left. \frac{\partial \ln Z[R, \tilde{R}]}{\partial \tilde{R}} \right|_R^* \Leftrightarrow R^* = \mathbf{J} * \langle \dot{n} \rangle,$$

$$0 = -N\tilde{R}^* + N \left. \frac{\partial \ln Z[R, \tilde{R}]}{\partial R} \right|_R^* \Leftrightarrow \tilde{R}^* = \langle \tilde{v} \rangle = 0. \quad (\text{C8})$$

Here,  $\langle \dot{n} \rangle(t)$  is the population-averaged firing rate. Inserting these saddle point solutions yields the partition functional, Eq. (20).

#### APPENDIX D: PERTURBATIVE EXPANSION

If fluctuations or nonlinearities are weak, a perturbative expansion around the mean-field theory can provide accurate estimates of fluctuation effects. For ease of notation, let  $\mathbf{x} = (v, \dot{n})^T$  and  $\tilde{\mathbf{x}} = (\tilde{v}, \tilde{n})^T$ . We expand  $\mathbf{x}$  around a background field,

$$\mathbf{x} = \bar{\mathbf{x}} + \delta\mathbf{x}, \quad \tilde{\mathbf{x}} = \tilde{\mathbf{x}}^* + \delta\tilde{\mathbf{x}}, \quad (\text{D1})$$

and collect terms up to linear order in the fluctuations in the free action  $S_0$ , with higher-order terms in the interacting part of the action  $S_V$ :

$$S = S_0 + S_V,$$

$$S_0 = \tilde{v}^T (\partial_t \bar{v} + \bar{v} + \bar{v} \bar{n} - E - \mathbf{J} * \langle \dot{n} \rangle) + \tilde{n}^T [\bar{n} - f(\bar{v})]$$

$$+ \tilde{v}^T (\partial_t + 1 + \bar{n}) \delta v + \tilde{v}^T \bar{v} \delta n + \tilde{n}^T \delta n - \tilde{n}^T f^{(1)} \delta v,$$

$$S_V = \tilde{v}^T \delta n \delta v - \sum_{p=2}^{\infty} \frac{\tilde{n}^p}{p!} f^{(p)}(\bar{v}) - \sum_{\substack{p,q=1 \\ p+q>2}}^{\infty} \frac{\tilde{n}^p}{p!} \frac{f^{(q)}}{q!} (\delta v)^q. \quad (\text{D2})$$

(We should also expand the response variable  $\tilde{\mathbf{x}}$  around a background field; we skip that here since in Appendix E we will constrain the background fields to be the mean trajectories, and the mean of response variables is 0.) A joint moment of  $v, \dot{n}$  is

$$\left\langle \prod_{i=1}^a \dot{n}(t_i) \prod_{j=1}^b v(t_j) \right\rangle$$

$$= \int \mathcal{D}\delta\mathbf{x} \int \mathcal{D}\tilde{\mathbf{x}} \prod_{i=1}^a \prod_{j=1}^b \dot{n}_i v_j \exp(-S_0 - S_V). \quad (\text{D3})$$

Expanding  $\exp(-S_V)$  in a functional Taylor series around a solution to the mean-field theory yields an expansion of the moment in terms of Gaussian integrals with respect to the free density  $\exp(-S_0)$ . Because of Wick's theorem, these integrals yield products of the propagators  $\bar{\Delta}$ . These expansions can be efficiently organized diagrammatically.

#### 1. Feynman rules

Here we give the Feynman rules for a perturbative expansion of statistics of the population-averaged system, Eq. (29), around the mean-field theory. This provides a

TABLE I. Vertices corresponding to the interacting action  $S_V$  in Eq. (D2).  $f^{(q)}$  is the  $q$ th derivative of the intensity function  $f$ , evaluated at the expansion point  $\bar{v}$ . The intensity function vertex  $f^{(q)}/q!$  also has the constraint that the sum of its in and out degrees must be at least three since the linear and bilinear terms in  $(\bar{n}, \bar{\delta}v)$  went in to the definitions of the background field and the propagators [Eq. (D2)].

Vertex	Factor	In degree ( $\delta v, \delta n$ )	Out degree ( $\bar{v}, \bar{n}$ )
•	$f(\bar{v})$	(0, 0)	(0, $\geq 2$ )
◦	$\frac{f^{(q)}}{q!} \delta(\sum_{i=1}^{K_{\text{in}}} \omega_i - \sum_{j=1}^{K_{\text{out}}} \omega_j)$	( $\geq 1, 0$ )	(0, $q$ ), $q \geq 1$
◦	$-\delta(\sum_{i=1}^{K_{\text{in}}} \omega_i - \sum_{j=1}^{K_{\text{out}}} \omega_j)$	(1, 1)	(1, 0)

graphical algorithm for computing arbitrary cumulants of the spike train  $\dot{n}$  or membrane voltage  $v$ . Moments can be composed from the cumulants by the appropriate Bell polynomials. We give the rules in the temporal frequency domain, for an expansion around a stationary point. Each cumulant can be decomposed into a sum of terms, each represented by a connected diagram. Those diagrams are composed of the vertices and edges in Tables I and II (Feynman diagrams generated with Ref. [137]).

The source vertex • emits factors of the response variable  $\bar{n}$  corresponding to spike fluctuations. Each internal vertex, ◦ receives configuration variables  $\delta n, \delta v$  and emits response variables  $\bar{n}, \bar{v}$ . In any connected diagram, each pair of vertices will be linked by at least one pair of configuration and response variables, e.g.,  $(\delta v, \bar{n})$ . Because of Wick's theorem, each pair of configuration and response variables is replaced by the corresponding propagator edge. For example, the pair  $\delta v, \bar{n}$  gives rise to the propagator  $\bar{\Delta}_{v, \bar{n}}$ . To calculate the joint cumulant  $\langle\langle \dot{n}^a v^b \rangle\rangle(\omega_1, \dots, \omega_{a+b-1})$ :

- (1) Place an external vertex for each of the  $a + b$  factors of  $\dot{n}$  and  $v$ .
- (2) Using the internal vertices and edges in Tables I and II, construct all connected graphs such that each external vertex has one incoming propagator edge. Each edge has its own frequency variable  $\omega_i$ .
- (3) To evaluate a diagram, multiply the factors of every edge and vertex together. Additionally, the sum of

TABLE II. Edges corresponding to the components of the propagator from  $S_0$  in Eq. (D2). Each measures the linear response of one configuration variable to a perturbation of another. For example,  $\bar{\Delta}_{v, \bar{n}}$  measures the linear response of the voltage to a spike fluctuation.

Edge	Propagator	Factor
————	$\bar{\Delta}_{n, \bar{n}}(\omega)$	$(1 + \bar{n} + i\omega)/(1 + \bar{n} + f^{(1)}\bar{v} + i\omega)$
~~~~~	$\bar{\Delta}_{v, \bar{n}}(\omega)$	$-\bar{v}/(1 + \bar{n} + f^{(1)}\bar{v} + i\omega)$
llllll	$\bar{\Delta}_{n, \bar{v}}(\omega)$	$f^{(1)}/(1 + \bar{n} + f^{(1)}\bar{v} + i\omega)$
-----	$\bar{\Delta}_{v, \bar{v}}(\omega)$	$1/(1 + \bar{n} + f^{(1)}\bar{v} + i\omega)$

external frequencies (those on the external vertices' incoming edges) is zero: also multiply by  $\delta(\sum_{i=1}^{a+b} \omega_i)$ . Finally, integrate over all of the internal frequencies: for each internal frequency  $\omega_i$ , integrate  $\int d\omega_i/2\pi$ .

- (4) Evaluate each connected diagram constructed in (2), and add the contributions of the diagrams.

We perform the integrals over internal frequencies analytically using the residue theorem. For a thorough introduction see, e.g., Refs. [135,138]. For an introduction to diagrammatic methods in the Poisson generalized linear model without resets (no self-coupling), see Ref. [55]. See Ref. [58] for detailed analytical calculations of the integrals over internal frequencies in that model.

Equations (12) and (24) are self-consistent one-loop equations of motion for the mean voltage and rate. The approximate joint cumulants appearing in them can also be calculated using the perturbative Feynman rules above; each is given by a tree diagram with the same two edges as in the loop diagram.

The corresponding perturbative corrections to the mean-field values of the voltage and rate are given by one-loop tadpole diagrams. For example, the perturbative one-loop correction to the mean-field voltage corresponding to

Eq. (14) is . The internal vertex ◦ in this

diagram carries a factor of  $-1$ . This corresponds to the sign this diagram appears with in Eqs. (12) and (24), opposite the sign in front of the second loop diagram. There is also a perturbative one-loop correction to the mean-field voltage

arising from the intensity vertex, . Both

nonlinearities also give rise to perturbative corrections to the mean-field rate. This is a difference with the self-consistent approach of the main text, where only one one-loop correction arises in each equation of motion (Appendix E).

## APPENDIX E: EFFECTIVE ACTION

Here we briefly derive the effective action. For a more detailed presentation see, e.g., Ref. [138], Chap. 7, or Ref. [135], Chaps. 11–14. For ease of notation, let  $\mathbf{x} = (v, \dot{n})^T$  and  $\tilde{\mathbf{x}} = (\bar{v}, \bar{n})^T$ . The cumulant-generating functional is

$$\exp W[\mathbf{j}, \tilde{\mathbf{j}}] = \int D\mathbf{x} \int D\tilde{\mathbf{x}} \exp \frac{1}{h} (-S[\mathbf{x}, \tilde{\mathbf{x}}] + \tilde{\mathbf{j}}^T \mathbf{x} + \mathbf{j}^T \tilde{\mathbf{x}}). \quad (\text{E1})$$

We have introduced a scale  $h$  into the exponent on the right-hand side. For physical calculations we will set  $h = 1$ . (Here,  $h$  has no relation to the  $h$  used in Sec. IX.) We

expand  $\mathbf{x}$  around a background field  $\bar{\mathbf{x}}$  [Eq. (D1)] and similarly for the response variable,  $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}^* + \delta\tilde{\mathbf{x}}$ . This yields

$$\exp\frac{1}{h}(hW[\mathbf{j},\tilde{\mathbf{j}}] - \tilde{\mathbf{j}}^T\bar{\mathbf{x}} - \mathbf{j}^T\tilde{\mathbf{x}}^*) = \int \mathcal{D}\delta\mathbf{x} \int \mathcal{D}\delta\tilde{\mathbf{x}} \exp\frac{1}{h}(-S + \tilde{\mathbf{j}}^T\delta\mathbf{x} + \mathbf{j}^T\delta\tilde{\mathbf{x}}). \quad (\text{E2})$$

We now require that our background field be the mean:  $\bar{\mathbf{x}} = \langle\mathbf{x}\rangle$  so that

$$0 = \langle\delta x_i\rangle = \frac{\partial}{\partial\tilde{j}_i} \exp\frac{1}{h}(hW[\mathbf{j},\tilde{\mathbf{j}}] - \tilde{\mathbf{j}}^T\bar{\mathbf{x}} - \mathbf{j}^T\tilde{\mathbf{x}}^*), \quad (\text{E3})$$

and similarly, we require  $\tilde{\mathbf{x}}^* = \langle\tilde{\mathbf{x}}\rangle$ . These requirements can be satisfied only at a stationary point of the exponent:

$$\begin{aligned} 0 &= \frac{\partial}{\partial j_i} (hW[\mathbf{j},\tilde{\mathbf{j}}] - \tilde{\mathbf{j}}^T\bar{\mathbf{x}} - \mathbf{j}^T\tilde{\mathbf{x}}^*), \\ 0 &= \frac{\partial}{\partial\tilde{j}_i} (hW[\mathbf{j},\tilde{\mathbf{j}}] - \tilde{\mathbf{j}}^T\bar{\mathbf{x}} - \mathbf{j}^T\tilde{\mathbf{x}}^*), \end{aligned} \quad (\text{E4})$$

which defines a Legendre transform from  $-hW$  to the effective action  $\Gamma$ :

$$\Gamma[\bar{\mathbf{x}},\tilde{\mathbf{x}}^*] = \sup_{\mathbf{j},\tilde{\mathbf{j}}} \tilde{\mathbf{j}}^T\bar{\mathbf{x}} + \mathbf{j}^T\tilde{\mathbf{x}}^* - hW[\mathbf{j},\tilde{\mathbf{j}}]. \quad (\text{E5})$$

Substituting the effective action in Eq. (E2) yields

$$\begin{aligned} &\exp\frac{1}{h}(-\Gamma[\bar{\mathbf{x}},\tilde{\mathbf{x}}^*] + S[\bar{\mathbf{x}},\tilde{\mathbf{x}}^*]) \\ &= \int \mathcal{D}\delta\mathbf{x} \int \mathcal{D}\delta\tilde{\mathbf{x}} \exp\frac{1}{h}(-S[\delta\mathbf{x},\delta\tilde{\mathbf{x}}] + \tilde{\mathbf{j}}^T\delta\mathbf{x} + \mathbf{j}^T\delta\tilde{\mathbf{x}}), \end{aligned} \quad (\text{E6})$$

where  $S[\bar{\mathbf{x}},\tilde{\mathbf{x}}^*]$  contains the terms in the  $S$  that depend only on  $\bar{\mathbf{x}},\tilde{\mathbf{x}}^*$  and not  $\delta\mathbf{x},\delta\tilde{\mathbf{x}}$ , and  $S[\delta\mathbf{x},\delta\tilde{\mathbf{x}}]$  contains the remaining terms of  $S$ . This has the form of a generating functional for  $\bar{\mathbf{x}},\tilde{\mathbf{x}}^*$ . The mean is a stationary point of  $\Gamma$ ; it obeys the equations of motion:

$$\begin{aligned} 0 &= \frac{\partial}{\partial\bar{x}_i} \Gamma, \\ 0 &= \frac{\partial}{\partial\tilde{x}_i^*} \Gamma. \end{aligned} \quad (\text{E7})$$

The loop expansion for the effective action is a diagrammatic equivalent of the saddle point expansion of the integrals over  $\delta\mathbf{x},\delta\tilde{\mathbf{x}}$  in Eq. (E6), without requiring that  $h$  be a bona fide small parameter [139,140]. The diagrams

TABLE III. Components of the propagator from  $S_0$  in Eq. (F2).

Edge	Propagator	Factor
————	$\bar{\Delta}_{n,\bar{n}}(\omega)$	$(1+i\omega)/(1+f^{(1)}r+i\omega)$
~~~~~	$\bar{\Delta}_{v,\bar{n}}(\omega)$	$-r/(1+f^{(1)}r+i\omega)$
~~~~~	$\bar{\Delta}_{n,\bar{v}}(\omega)$	$f^{(1)}/(1+f^{(1)}r+i\omega)$
-----	$\bar{\Delta}_{v,\bar{v}}(\omega)$	$1/(1+f^{(1)}r+i\omega)$

contributing to the equations of motion for  $\bar{\mathbf{x}},\tilde{\mathbf{x}}^*$  are one-line-irreducible vacuum diagrams (those that cannot be disconnected by cutting one edge; see, e.g., Ref. [135], Chaps. 11.4, 13.3). Only those diagrams with a vertex carrying the appropriate factor of  $\tilde{x}_i^*$  will contribute to the equation of motion  $0 = (\partial/\partial\tilde{x}_i^*)\Gamma$  (and similarly for  $\partial/\partial\bar{x}_i$ ), which is why the one-loop equations of motion each have only one loop correction.

## APPENDIX F: CONNECTIVITY AVERAGE FOR THE LINEAR-RESET MODEL

The connectivity-averaged action for the linear-reset model is

$$\begin{aligned} S[v,\dot{n},\bar{v},\bar{n}] &= \bar{v}^T(\partial_t v + v + r\dot{n} - E - J\langle\dot{n}\rangle) \\ &\quad + \bar{n}^T\dot{n} - (e^{\bar{n}} - 1)^T f(v), \end{aligned} \quad (\text{F1})$$

and expanding around a solution to the mean-field theory  $\partial_t\bar{v} = -\bar{v} + (J-r)\langle\dot{n}\rangle + E$  yields the free and interacting actions:

$$\begin{aligned} S_0 &= \bar{v}^T(\partial_t\bar{v} + \bar{v} + r\bar{n} - E - J*\langle\dot{n}\rangle) + \bar{n}^T[\bar{n} - f(\bar{v})] \\ &\quad + \bar{v}^T(\partial_t + 1)\delta v + \bar{v}^T r\delta n + \bar{n}^T\delta n - \bar{n}^T f^{(1)}\delta v, \\ S_V &= -\sum_{p=2}^{\infty} \frac{\bar{n}^p}{p!} \bar{f} - \sum_{\substack{p,q=1 \\ p+q>2}}^{\infty} \frac{\bar{n}^p}{p!} \frac{f^{(q)}}{q!} (\delta v)^q. \end{aligned} \quad (\text{F2})$$

The components of the propagator for this model are given in Table III. It has the same source and intensity vertices as the stochastic LIF model (the first two entries in Table I), but lacks the reset vertex.

- 
- [1] W. J. Freeman, *Mass Action in the Nervous System: Examination of the Neurophysiological Basis of Adaptive Behavior Through the EEG* (Academic Press, New York, 1975).
  - [2] S. Coombes, *Large-scale neural dynamics: Simple and complex*, *NeuroImage* **52**, 731 (2010).
  - [3] P. C. Bressloff, *Spatiotemporal dynamics of continuum neural fields*, *J. Phys. A* **45**, 033001 (2011).
  - [4] R. Moran, D. Pinotsis, and K. Friston, *Neural masses and fields in dynamic causal modeling*, *Front. Comput. Neurosci.* **7**, 57 (2013).

- [5] S. Grossberg, *Learning and energy-entropy dependence in some nonlinear functional-differential systems*, *Bull. Am. Math. Soc.* **75**, 1238 (1969).
- [6] S.-I. Amari, *Characteristics of randomly connected threshold-element networks and network systems*, *Proc. IEEE* **59**, 35 (1971).
- [7] S.-I. Amari, *Characteristics of random nets of analog neuron-like elements*, *IEEE Trans. Syst. Man. Cybernet. SMC-2*, 643 (1972).
- [8] H. R. Wilson and J. D. Cowan, *Excitatory and inhibitory interactions in localized populations of model neurons*, *Biophys. J.* **12**, 1 (1972).
- [9] H. R. Wilson and J. D. Cowan, *A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue*, *Kybernetik* **13**, 55 (1973).
- [10] I. Ginzburg and H. Sompolinsky, *Theory of correlations in stochastic neural networks*, *Phys. Rev. E* **50**, 3171 (1994).
- [11] T. Ohira and J. D. Cowan, *Master-equation approach to stochastic neurodynamics*, *Phys. Rev. E* **48**, 2259 (1993).
- [12] M. A. Buice and J. D. Cowan, *Field-theoretic approach to fluctuation effects in neural networks*, *Phys. Rev. E* **75**, 051919 (2007).
- [13] P. C. Bressloff, *Stochastic neural field theory and the system-size expansion*, *SIAM J. Appl. Math.* **70**, 1488 (2010).
- [14] A. L. Hodgkin and A. F. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, *J. Physiol.* **117**, 500 (1952).
- [15] C. C. Chow and Y. Karimipناه, *Before and beyond the Wilson–Cowan equations*, *J. Neurophys.* **123**, 1645 (2020).
- [16] N. Brunel, V. Hakim, and M. J. Richardson, *Single neuron dynamics and computation*, *Curr. Opin. Neurobiol.* **25**, 149 (2014).
- [17] B. Doiron, A. Litwin-Kumar, R. Rosenbaum, G. K. Ocker, and K. Josić, *The mechanics of state-dependent neural correlations*, *Nat. Neurosci.* **19**, 383 (2016).
- [18] B. W. Knight, *Dynamics of encoding in a population of neurons*, *J. Gen. Physiol.* **59**, 734 (1972).
- [19] L. M. Ricciardi, *Diffusion Processes and Related Topics in Biology* (Springer, New York, 1977).
- [20] D. J. Amit and N. Brunel, *Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex*, *Cereb. Cortex* **7**, 237 (1997).
- [21] N. Brunel, *Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons*, *J. Comput. Neurosci.* **8**, 183 (2000).
- [22] B. Lindner and L. Schimansky-Geier, *Transmission of noise coded versus additive signals through a neuronal ensemble*, *Phys. Rev. Lett.* **86**, 2934 (2001).
- [23] B. Doiron, B. Lindner, A. Longtin, L. Maler, and J. Bastian, *Oscillatory activity in electrosensory neurons increases with the spatial correlation of the stochastic input stimulus*, *Phys. Rev. Lett.* **93**, 048101 (2004).
- [24] B. Lindner, B. Doiron, and A. Longtin, *Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback*, *Phys. Rev. E* **72**, 061919 (2005).
- [25] B. Lindner, *Superposition of many independent spike trains is generally not a Poisson process*, *Phys. Rev. E* **73**, 022901 (2006).
- [26] R. Moreno-Bote and N. Parga, *Auto- and crosscorrelograms for the spike response of leaky integrate-and-fire neurons with slow synapses*, *Phys. Rev. Lett.* **96**, 028101 (2006).
- [27] T. Schwalger, F. Droste, and B. Lindner, *Statistical structure of neural spiking under non-Poissonian or other non-white stimulation*, *J. Comput. Neurosci.* **39**, 29 (2015).
- [28] S. Vellmer and B. Lindner, *Theory of spike-train power spectra for multidimensional integrate-and-fire neurons*, *Phys. Rev. Res.* **1**, 023024 (2019).
- [29] W. Gerstner, *Time structure of the activity in neural network models*, *Phys. Rev. E* **51**, 738 (1995).
- [30] W. Gerstner, *Population dynamics of spiking neurons: Fast transients, asynchronous states, and locking*, *Neural Comput.* **12**, 43 (2000).
- [31] C. Meyer and C. v. Vreeswijk, *Temporal correlations in stochastic networks of spiking neurons*, *Neural Comput.* **14**, 369 (2002).
- [32] M. Deger, T. Schwalger, R. Naud, and W. Gerstner, *Fluctuations and information filtering in coupled populations of spiking neurons with adaptation*, *Phys. Rev. E* **90**, 062704 (2014).
- [33] G. Dumont, A. Payeur, and A. Longtin, *A stochastic-field description of finite-size spiking neural networks*, *PLoS Comput. Biol.* **13**, e1005691 (2017).
- [34] T. Schwalger, M. Deger, and W. Gerstner, *Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size*, *PLoS Comput. Biol.* **13**, e1005507 (2017).
- [35] M. Mattia and P. Del Giudice, *Population dynamics of interacting spiking neurons*, *Phys. Rev. E* **66**, 051917 (2002).
- [36] B. Pietras, N. Gallice, and T. Schwalger, *Low-dimensional firing-rate dynamics for populations of renewal-type spiking neurons*, *Phys. Rev. E* **102**, 022407 (2020).
- [37] P. C. Martin, E. D. Siggia, and H. A. Rose, *Statistical dynamics of classical systems*, *Phys. Rev. A* **8**, 423 (1973).
- [38] C. D. Dominicis, *Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques*, *J. Phys. (Paris), Colloq.* **37**, C1-247 (1976).
- [39] H.-K. Janssen, *On a Lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties*, *Z. Phys. B* **23**, 377 (1976).
- [40] R. V. Jensen, *Functional integral approach to classical statistical dynamics*, *J. Stat. Phys.* **25**, 183 (1981).
- [41] H. Sompolinsky, A. Crisanti, and H. J. Sommers, *Chaos in random neural networks*, *Phys. Rev. Lett.* **61**, 259 (1988).
- [42] M. Stern, H. Sompolinsky, and L. F. Abbott, *Dynamics of random neural networks with bistable units*, *Phys. Rev. E* **90**, 062710 (2014).
- [43] P. C. Bressloff, *Path-integral methods for analyzing the effects of fluctuations in stochastic hybrid neural networks*, *J. Math. Neurosci.* **5**, 4 (2015).
- [44] A. van Meegen, T. Kühn, and M. Helias, *Large-deviation approach to random recurrent neuronal networks: Parameter inference and fluctuation-induced transitions*, *Phys. Rev. Lett.* **127**, 158302 (2021).
- [45] U. Pereira and N. Brunel, *Attractor dynamics in networks with learning rules inferred from in vivo data*, *Neuron* **99**, 227 (2018).

- [46] J. Schuecker, S. Goedeke, and M. Helias, *Optimal sequence memory in driven random networks*, *Phys. Rev. X* **8**, 041029 (2018).
- [47] J. Aljadeff, M. Stern, and T. Sharpee, *Transition to chaos in random networks with cell-type-specific connectivity*, *Phys. Rev. Lett.* **114**, 088101 (2015).
- [48] F. Mastrogiuseppe and S. Ostojic, *Linking connectivity, dynamics, and computations in low-rank recurrent neural networks*, *Neuron* **99**, 609 (2018).
- [49] F. Schuessler, A. Dubreuil, F. Mastrogiuseppe, S. Ostojic, and O. Barak, *Dynamics of random recurrent networks with correlated low-rank structure*, *Phys. Rev. Res.* **2**, 013111 (2020).
- [50] I. D. Landau and H. Sompolinsky, *Macroscopic fluctuations emerge in balanced networks with incomplete recurrent alignment*, *Phys. Rev. Res.* **3**, 023171 (2021).
- [51] M. A. Buice and C. C. Chow, *Dynamic finite size effects in spiking neural networks*, *PLoS Comput. Biol.* **9**, e1002872 (2013).
- [52] J. Kadmon and H. Sompolinsky, *Transition to chaos in random neuronal networks*, *Phys. Rev. X* **5**, 041030 (2015).
- [53] O. Harish and D. Hansel, *Asynchronous rate chaos in spiking neuronal circuits*, *PLoS Comput. Biol.* **11**, e1004266 (2015).
- [54] F. Mastrogiuseppe and S. Ostojic, *Intrinsically-generated fluctuating activity in excitatory-inhibitory networks*, *PLoS Comput. Biol.* **13**, e1005498 (2017).
- [55] G. K. Ocker, K. Josić, E. Shea-Brown, and M. A. Buice, *Linking structure and activity in nonlinear spiking networks*, *PLoS Comput. Biol.* **13**, e1005583 (2017).
- [56] B. A. W. Brinkman, F. Rieke, E. Shea-Brown, and M. A. Buice, *Predicting how and when hidden neurons skew measured synaptic interactions*, *PLoS Comput. Biol.* **14**, e1006490 (2018).
- [57] D. Todorov and W. Truccolo, *Stability of stochastic finite-size spiking-neuron networks: Comparing mean-field, 1-loop correction and quasi-renewal approximations*, in *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, New York, 2019), pp. 4380–4386, ISSN: 1558-4615.
- [58] M. Kordovan and S. Rotter, *Spike train cumulants for linear-nonlinear Poisson cascade models*, [arXiv:2001.05057](https://arxiv.org/abs/2001.05057).
- [59] J. S. Griffith, *On the stability of brain-like structures*, *Biophys. J.* **3**, 299 (1963).
- [60] M. V. Tsodyks, W. E. Skaggs, T. J. Sejnowski, and B. L. McNaughton, *Paradoxical effects of external modulation of inhibitory interneurons*, *J. Neurosci.* **17**, 4382 (1997).
- [61] H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, and D. Ferster, *Inhibitory stabilization of the cortical network underlies visual surround suppression*, *Neuron* **62**, 578 (2009).
- [62] H. K. Kato, S. K. Asinof, and J. S. Isaacson, *Network-level control of frequency tuning in auditory cortex*, *Neuron* **95**, 412 (2017).
- [63] H. Adesnik, *Synaptic mechanisms of feature coding in the visual cortex of awake mice*, *Neuron* **95**, 1147 (2017).
- [64] A. Sanzeni, B. Akitake, H. C. Goldbach, C. E. Leedy, N. Brunel, and M. H. Histed, *Inhibition stabilization is a widespread property of cortical networks*, *eLife* **9**, e54875 (2020).
- [65] Z. Zhou, Y. Zhang, Y. Xie, Z. Wang, and J.-G. Liu, *Investigating the integrate and fire model as the limit of a random discharge model: A stochastic analysis perspective*, *Math. Neurosci. Appl.* **1** (2021) 10.46298/mna.7203.
- [66] D. J. Pinto, J. C. Brumberg, D. J. Simons, G. B. Ermentrout, and R. Traub, *A quantitative population model of whisker barrels: Re-examining the Wilson-Cowan equations*, *J. Comput. Neurosci.* **3**, 247 (1996).
- [67] K. D. Miller and T. W. Troyer, *Neural noise can explain expansive, power-law nonlinearities in neural response functions, power-law nonlinearities in neural response functions*, *J. Neurophysiol.* **87**, 653 (2002).
- [68] D. Hansel and C. v. Vreeswijk, *How noise contributes to contrast invariance of orientation tuning in cat visual cortex*, *J. Neurosci.* **22**, 5118 (2002).
- [69] N. J. Priebe, F. Mechler, M. Carandini, and D. Ferster, *The contribution of spike threshold to the dichotomy of cortical simple and complex cells*, *Nat. Neurosci.* **7**, 1113 (2004).
- [70] N. J. Priebe and D. Ferster, *Mechanisms underlying cross-orientation suppression in cat visual cortex*, *Nat. Neurosci.* **9**, 552 (2006).
- [71] D. Linaro, G. K. Ocker, B. Doiron, and M. Giugliano, *Correlation transfer by layer 5 cortical neurons under recreated synaptic inputs in vitro*, *J. Neurosci.* **39**, 7648 (2019).
- [72] W. Gerstner and J. L. van Hemmen, *Universality in neural networks: The importance of the ‘mean firing rate’*, *Biol. Cybern.* **67**, 195 (1992).
- [73] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, Cambridge, England, 2014).
- [74] D. Cox and V. Isham, *Point Processes*, Monographs on Statistics and Applied Probability (CRC Press, Boca Raton, FL, 1980).
- [75] P. Robert and J. Touboul, *On the dynamics of random neuronal networks*, *J. Stat. Phys.* **165**, 545 (2016).
- [76] B. Rudy, G. Fishell, S. Lee, and J. Hjerling-Leffler, *Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons*, *Dev. Neurobiol.* **71**, 45 (2011).
- [77] B. Tasic *et al.*, *Adult mouse cortical cell taxonomy revealed by single cell transcriptomics*, *Nat. Neurosci.* **19**, 335 (2016).
- [78] Z. Yao *et al.*, *A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation*, *Cell* **184**, 3222 (2021).
- [79] E. M. Callaway *et al.*, *A multimodal cell census and atlas of the mammalian primary motor cortex*, *Nature (London)* **598**, 86 (2021).
- [80] C. K. Pfeffer, M. Xue, M. He, Z. J. Huang, and M. Scanziani, *Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons*, *Nat. Neurosci.* **16**, 1068 (2013).
- [81] R. Tremblay, S. Lee, and B. Rudy, *GABAergic Interneurons in the neocortex: From cellular properties to Circuits*, *Neuron* **91**, 260 (2016).

- [82] S. C. Seeman *et al.*, *Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex*, *eLife* **7**, e37349 (2018).
- [83] T. A. Hage, A. Bosma-Moody, C. A. Baker, M. B. Kratz, L. Campagnola, T. Jarsky, H. Zeng, and G. J. Murphy, *Synaptic connectivity to L2/3 of primary visual cortex measured by two-photon optogenetic stimulation*, *eLife* **11**, e71103 (2022).
- [84] M. J. Chacron and J. Bastian, *Population coding by electrosensory neurons*, *J. Neurophys.* **99**, 1825 (2008).
- [85] J. Yu and D. Ferster, *Membrane potential synchrony in primary visual cortex during sensory stimulation*, *Neuron* **68**, 1187 (2010).
- [86] S. Giridhar, B. Doiron, and N. N. Urban, *Timescale-dependent shaping of correlation by olfactory bulb lateral inhibition*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5843 (2011).
- [87] R. Kempter, W. Gerstner, and J. L. van Hemmen, *Hebbian learning and spiking neurons*, *Phys. Rev. E* **59**, 4498 (1999).
- [88] M. Gilson, A. Burkitt, and J. L. van Hemmen, *STDP in recurrent neuronal networks*, *Front. Comput. Neurosci.* **4**, 23 (2010).
- [89] B. Babadi and L. F. Abbott, *Pairwise analysis can account for network structures arising from spike-timing dependent plasticity*, *PLoS Comput. Biol.* **9**, e1002906 (2013).
- [90] G. K. Ocker, A. Litwin-Kumar, and B. Doiron, *Self-organization of microcircuits in networks of spiking neurons with plastic synapses*, *PLoS Comput. Biol.* **11**, e1004458 (2015).
- [91] N. R. Tannenbaum and Y. Burak, *Shaping neural circuits by high order synaptic interactions*, *PLoS Comput. Biol.* **12**, e1005056 (2016).
- [92] G. K. Ocker and B. Doiron, *Training and spontaneous reinforcement of neuronal assemblies by spike timing plasticity*, *Cereb. Cortex* **29**, 937 (2018).
- [93] L. Montangie, C. Miehl, and J. Gjorgjieva, *Autonomous emergence of connectivity assemblies via spike triplet interactions*, *PLoS Comput. Biol.* **16**, e1007835 (2020).
- [94] D. R. Cox, *Renewal Theory* (Methuen, London, 1962).
- [95] A. Litwin-Kumar, R. Rosenbaum, and B. Doiron, *Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes*, *J. Neurophys.* **115**, 1399 (2016).
- [96] T. Ohyama *et al.*, *A multilevel multimodal circuit enhances action selection in Drosophila*, *Nature (London)* **520**, 633 (2015).
- [97] M. E. Berck *et al.*, *The wiring diagram of a glomerular olfactory system*, *eLife* **5**, e14859 (2016).10.7554/eLife.14859
- [98] A. A. Wanner, C. Genoud, T. Masudi, L. Siksou, and R. W. Friedrich, *Dense EM-based reconstruction of the interglomerular projectome in the zebrafish olfactory bulb*, *Nat. Neurosci.* **19**, 816 (2016).
- [99] J. L. Morgan, D. R. Berger, A. W. Wetzel, and J. W. Lichtman, *The fuzzy logic of network connectivity in mouse visual thalamus*, *Cell* **165**, 192 (2016).
- [100] I. Larderet, P. M. Fritsch, N. Gendre, G. L. Neagu-Maier, R. D. Fetter, C. M. Schneider-Mizell, J. W. Truman, M. Zlatic, A. Cardona, and S. G. Sprecher, *Organization of the Drosophila larval visual circuit*, *eLife* **6**, e28387 (2017).
- [101] D. G. C. Hildebrand *et al.*, *Whole-brain serial-section electron microscopy in larval zebrafish*, *Nature (London)* **545**, 345 (2017).
- [102] K. Eichler *et al.*, *The complete connectome of a learning and memory centre in an insect brain*, *Nature (London)* **548**, 175 (2017).
- [103] Z. Zheng *et al.*, *A complete electron microscopy volume of the brain of adult Drosophila melanogaster*, *Cell* **174**, 730 (2018).
- [104] A. Motta, M. Berning, K. M. Boergens, B. Staffler, M. Beining, S. Loomba, C. Schramm, P. Hennig, H. Wissler, and M. Helmstaedter, *Dense connectomic reconstruction in layer 4 of the somatosensory cortex*, *Science* **366**, eaay3134 (2019).
- [105] S. J. Cook *et al.*, *Whole-animal connectomes of both Caenorhabditis elegans sexes*, *Nature (London)* **571**, 63 (2019).
- [106] L. K. Scheffer *et al.*, *A connectome of the adult Drosophila central brain*, *eLife* **9**, e57433 (2020).
- [107] J. A. Bae *et al.* (MICrONS Consortium), *Functional connectomics spanning multiple areas of mouse visual cortex*, *bioRxiv*, 10.1101/2021.07.28.454025.
- [108] A. Vishwanathan *et al.*, *Predicting modular functions and neural coding of behavior from a synaptic wiring diagram*, *bioRxiv*, 10.1101/2020.10.28.359620.
- [109] C. Curto, J. Geneson, and K. Morrison, *Fixed points of competitive threshold-linear networks*, *Neural Comput.* **31**, 94 (2019).
- [110] C. Curto, C. Langdon, and K. Morrison, *Combinatorial geometry of threshold-linear networks*, *arXiv:2008.01032*.
- [111] D. E. Santander, S. Ebli, A. Patania, N. Sanderson, F. Burtscher, K. Morrison, and C. Curto, *Nerve theorems for fixed points of neural networks*, in *Research in Computational Topology 2*, Association for Women in Mathematics Series, edited by E. Gasparovic, V. Robins, and K. Turner (Springer International Publishing, Cham, 2022), pp. 129–165.
- [112] C. Parmelee, J. L. Alvarez, C. Curto, and K. Morrison, *Sequential attractors in combinatorial threshold-linear networks*, *SIAM J. Appl. Dyn. Syst.* **21**, 1597 (2022).
- [113] A. G. Hawkes, *Spectra of some self-exciting and mutually exciting point processes*, *Biometrika* **58**, 83 (1971).
- [114] T. J. Sejnowski, *On the stochastic dynamics of neuronal interaction*, *Biol. Cybern.* **22**, 203 (1976).
- [115] S. Ostojic, N. Brunel, and V. Hakim, *How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains*, *J. Neurosci.* **29**, 10234 (2009).
- [116] V. Pernice, B. Staude, S. Cardanobile, and S. Rotter, *How structure determines correlations in neuronal networks*, *PLoS Comput. Biol.* **7**, e1002059 (2011).
- [117] J. Trousdale, Y. Hu, E. Shea-Brown, and K. Josić, *Impact of network structure and cellular response on spike time correlations*, *PLoS Comput. Biol.* **8**, e1002408 (2012).
- [118] S. Jovanović and S. Rotter, *Interplay between graph topology and correlations of third order in spiking neuronal networks*, *PLoS Comput. Biol.* **12**, e1004963 (2016).
- [119] M. J. Cáceres, J. A. Carrillo, and B. Perthame, *Analysis of nonlinear noisy integrate & fire neuron models: Blow-up and steady states*, *J. Math. Neurosci.* **1**, 7 (2011).

- [120] C. R. Laing and C. C. Chow, *Stationary bumps in networks of spiking neurons*, *Neural Comput.* **13**, 1473 (2001).
- [121] R. Rosenbaum and B. Doiron, *Balanced networks of spiking neurons with spatially dependent recurrent connections*, *Phys. Rev. X* **4**, 021039 (2014).
- [122] A. De Masi, A. Galves, E. Löcherbach, and E. Presutti, *Hydrodynamic limit for interacting neurons*, *J. Stat. Phys.* **158**, 866 (2015).
- [123] N. Fournier and E. Löcherbach, *On a toy model of interacting neurons*, *Ann. Inst. Henri Poincaré* **52**, 1844 (2016).
- [124] A. Duarte and G. Ost, *A model for neural activity in the absence of external stimuli*, [arXiv:1410.6086](https://arxiv.org/abs/1410.6086).
- [125] Q. Cormier, E. Tanré, and R. Veltz, *Long time behavior of a mean-field model of interacting neurons*, *Stoch. Proc. Appl.* **130**, 2553 (2020).
- [126] V. Schmutz, E. Löcherbach, and T. Schwalger, *On a finite-size neuronal population equation*, [arXiv:2106.14721](https://arxiv.org/abs/2106.14721).
- [127] F. Baccelli and T. Taillefumier, *Replica-mean-field limits for intensity-based neural networks*, *SIAM J. Appl. Dyn. Syst.* **18**, 1756 (2019).
- [128] F. Baccelli and T. Taillefumier, *The pair-replica-mean-field limit for intensity-based neural networks*, *SIAM J. Appl. Dyn. Syst.* **20**, 165 (2021).
- [129] L. Yu and T. O. Taillefumier, *Metastable spiking networks in the replica-mean-field limit*, *PLoS Comput. Biol.* **18**, e1010215 (2022).
- [130] G. K. Ocker, *Dynamics of stochastic integrate-and-fire networks*, *Phys. Rev. X* **12**, 041007 (2022).
- [131] G. K. Ocker, *Retraction: Dynamics of stochastic integrate-and-fire networks* [*Phys. Rev. X* **12**, 041007 (2022)]; **13**, 029904 (2023).
- [132] G. K. Ocker, [https://github.com/gkocker/lif\\_stochastic](https://github.com/gkocker/lif_stochastic).
- [133] C. C. Chow and M. A. Buice, *Path integral methods for stochastic differential equations*, *J. Math. Neurosci.* **5**, 8 (2015).
- [134] J. A. Hertz, Y. Roudi, and P. Sollich, *Path integral methods for the dynamics of stochastic and disordered systems*, *J. Phys. A* **50**, 033001 (2016).
- [135] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks*, Lecture Notes in Physics (Springer International Publishing, Cham, 2020).
- [136] A. Galves and E. Löcherbach, *Infinite systems of interacting chains with memory of variable length—A stochastic model for biological neural nets*, *J. Stat. Phys.* **151**, 896 (2013).
- [137] J. P. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, *Comput. Phys. Commun.* **210**, 103 (2017).
- [138] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon Press, Oxford, 2002).
- [139] S. Coleman and E. Weinberg, *Radiative corrections as the origin of spontaneous symmetry breaking*, *Phys. Rev. D* **7**, 1888 (1973).
- [140] M. A. Buice, J. D. Cowan, and C. C. Chow, *Systematic fluctuation expansion for neural network activity equations*, *Neural Comput.* **22**, 377 (2010).