Hopf Bifurcation in Mean Field Explains Critical Avalanches in Excitation-Inhibition Balanced Neuronal Networks: A Mechanism for Multiscale Variability

Junhao Liang\(^1,2\), Tianshou Zhou\(^2\) and Changsong Zhou\(^1,3\)*

\(^1\) Department of Physics, Centre for Nonlinear Studies, Beijing-Hong Kong-Singapore Joint Centre for Nonlinear and Complex Systems, Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Kowloon Tong, Hong Kong, \(^2\) Key Laboratory of Computational Mathematics, Guangdong Province, and School of Mathematics, Sun Yat-sen University, Guangzhou, China, \(^3\) Department of Physics, Zhejiang University, Hangzhou, China

Cortical neural circuits display highly irregular spiking in individual neurons but variably sized collective firing, oscillations and critical avalanches at the population level, all of which have functional importance for information processing. Theoretically, the balance of excitation and inhibition inputs is thought to account for spiking irregularity and critical avalanches may originate from an underlying phase transition. However, the theoretical reconciliation of these multilevel dynamic aspects in neural circuits remains an open question. Herein, we study excitation-inhibition (E-I) balanced neuronal network with biologically realistic synaptic kinetics. It can maintain irregular spiking dynamics with different levels of synchrony and critical avalanches emerge near the synchronous transition point. We propose a novel semi-analytical mean-field theory to derive the field equations governing the network macroscopic dynamics. It reveals that the E-I balanced state of the network manifesting irregular individual spiking is characterized by a macroscopic stable state, which can be either a fixed point or a periodic motion and the transition is predicted by a Hopf bifurcation in the macroscopic field. Furthermore, by analyzing public data, we find the coexistence of irregular spiking and critical avalanches in the spontaneous spiking activities of mouse cortical slice in vitro, indicating the universality of the observed phenomena. Our theory unveils the mechanism that permits complex neural activities in different spatiotemporal scales to coexist and elucidates a possible origin of the criticality of neural systems. It also provides a novel tool for analyzing the macroscopic dynamics of E-I balanced networks and its relationship to the microscopic counterparts, which can be useful for large-scale modeling and computation of cortical dynamics.

Keywords: excitation-inhibition balance, neuronal avalanches, criticality, cortical dynamics, mean-field theory
INTRODUCTION

The mammalian brain consists of tens of billions of neurons, which process information and communicate through electrophysiological action potentials, also known as spikes. This large number of neurons exhibit diverse spiking behaviors across broad ranges of spatial and temporal scales. Understanding the origin and dynamic mechanism of this complexity is crucial for the advancement of neurobiology, the development of therapies for brain diseases and the future design of brain-inspired intelligent systems.

Two striking features can be simultaneously observed at different levels of cortical neuronal systems: 1) irregularity in spiking times, indicated by seemingly random spiking time that resembles Poisson process (Softky and Koch, 1993; Holt et al., 1996) of individual neurons; 2) variability in population firing rates, manifested in widely observed collective neural activities such as population oscillations (Brunel and Wang, 2003; Herrmann et al., 2004) and critical neural avalanches (Beggs and Plenz, 2003; Gireesh and Plenz, 2008; Bellay et al., 2015; Shew et al., 2015; Ma et al., 2019), etc. Biologically, the spiking irregularity has been proposed to originate from the balance between excitation (E) and inhibition (I) inputs so that spiking of neurons is driven by fluctuations (Shu et al., 2003; Okun and Lampl, 2008; Xue et al., 2014), and has been associated with functional advantages in efficient coding and information processing (Denève and Machens, 2016). The emergence of collective cortical activities originates in the fact that neurons interact through recurrent networks (Abeles, 1991), in which dynamic activities can reverberate. As a result, dynamic correlations arise from structural correlations. In particular, even weak pairwise correlation is sufficient to induce strongly correlated collective network activities (Schneidman et al., 2006). Collective neural activities can emerge with different amplitudes and are often organized as critical avalanches with various sizes. These avalanches are cascades of activity bursts in neuronal networks. At criticality, the size and duration of avalanches are approximately distributed according to power-laws, with critical exponents satisfying the scaling relation (Friedman et al., 2012; Fontenele et al., 2019). Avalanches in the critical state can maximize the informational complexity and variability, and are thought to have functional advantages in information processing (Kinouchi and Copelli, 2006; Shew et al., 2009, 2011).

Traditional mean-field theory of E-I balanced networks ( VAN Vreeswijk and Sompolinsky, 1996, 1998; Renart et al., 2010) with binary neuron and instantaneous synapse explains the spiking irregularity of individual neurons. However, it fails to account for collective neural activities, because it predicts an asynchronous dynamic state with vanishing correlation in unstructured (i.e., random topology) networks. Such vanishing correlation arises due to sparse network connectivity (Van Vreeswijk and Sompolinsky, 1996) or shared excitatory and inhibitory inputs cancelling correlation in dense recurrent networks (Renart et al., 2010). In terms of rate coding, the asynchronous state is not efficient for information processing, as the population firing rate only exhibits a linear response to the input rate (Van Vreeswijk and Sompolinsky, 1996) but without firing rate variability on faster time scales. In this case, the whole network acts as a rate unit for computation. The traditional E-I balanced theory can be generalized in two directions, which are biologically more plausible: structured networks and synaptic kinetics. Firstly, heterogeneous neural network structures (Landau et al., 2016) can induce firing rate variability. For example, clustered network structures (Litwin-Kumar and Doiron, 2012) can induce slow firing rate oscillations and show stimulus-induced variability reductions. Hierarchical modular networks (Wang et al., 2011) can support self-sustained firing rate oscillations across different levels. Spatial networks involving a distance-dependent coupling rule can unveil the distance-tuned correlation relation (Rosenbaum et al., 2017; Darshan et al., 2018) and the emergence of propagating waves (Keane and Gong, 2015; Keane et al., 2018; Gu et al., 2019; Huang et al., 2019) observed in experiments. Secondly, even in unstructured networks, network firing rate oscillations can be induced by realistic synaptic filtering kinetics (Brunel and Wang, 2003; Yang et al., 2017). Such oscillations typically occur in cases where the synaptic decay time scales of inhibition are slower than excitation, which is actually a biologically plausible situation if the synaptic receptors under consideration are AMPA for excitation and GABA for inhibition (Salin and Prince, 1996; Zhou and Hablitz, 1998). More importantly, network oscillation can be sparsely participated by subgroups of neurons, thus preserving the irregular spiking feature of individual neurons (Brunel, 2000; Brunel and Hakim, 2008). Note, however, that theoretical analysis of the macroscopic dynamics of E-I neural network with synaptic kinetics is very difficult. Existing theory is very limited, e.g., by requiring very special assumptions, such as the Lorenz distribution of certain parameters (Dumont and Gutkin, 2019).

Critical avalanches can also rise from neural dynamics under unstructured network topology (Beggs and Plenz, 2003; Kinouchi and Copelli, 2006), while its dynamic origin is still controversial. Previous theories have suggested that critical neural avalanches may arise at the edge of a phase transition. Early studies indicated that it may occur between a quiescent and active phase from critical branching processes (Beggs and Plenz, 2003; Haldeman and Beggs, 2005), while later experimental (Fontenele et al., 2019) and theoretical (di Santo et al., 2018; Dalla Porta and Copelli, 2019) studies also proposed that it may occur near the onset of synchrony. Mechanisms other than criticality that generate avalanches with power-law distribution have also been proposed (Martinnello et al., 2017; Touboul and Destexhe, 2017; Zilting and Priesemann, 2019b). The emergence of critical avalanches has also been proposed to be closely related to the maintenance of E-I balance (Lombardi et al., 2012; Poil et al., 2012; Yang et al., 2012; Gautam et al., 2015). Nevertheless, the exact relationship between neural criticality and E-I balance remains poorly understood, as previous modeling studies of critical avalanches often did not consider the properties of individual neuronal spiking.

Here, we try to address the above important open questions. In particular, how E-I balance induced irregular spiking reconciles with collective neural activities? E-I neural networks can organize into a ‘sparse synchrony’ state (Brunel, 2000; Brunel and Hakim, 2008) where neurons are remained fluctuation-driven to spike irregularly whereas firing rate oscillation emerges in population...
level. However, the relationship between E-I balance and sparse synchrony is less clear. Most importantly, the mechanism by which E-I balance induced irregular spiking can coexist with critical neural activities (Bellay et al., 2015) in recurrent neural circuit remains unclear. In this work, we first re-examine the dynamics of integrate-and-fire (IF) E-I neuronal network with realistic synaptic kinetics, which can manifest individually irregular spiking activities with different synchronous levels. The network firing rate dynamics can be effectively captured by a set of macroscopic field equations derived by a novel semi-analytical mean-field theory. An advantage of our theory is that it is simple and does not require special properties of the model. The synchronous transition point where network firing rate oscillation emerges is predicted by a Hopf bifurcation in the field equations. We find that critical macroscopic avalanche dynamics emerges near the onset of synchronization, with critical exponents approximately satisfying the scaling relations, which manifests the hallmark of criticality (Sethna et al., 2001). The mechanism of critical avalanches could be understood as demographic noise-driven random walks near a macroscopic bifurcation point. On this basis, we propose that the E-I balanced state in the microscopic spiking network corresponds to a stable macroscopic state in the field equations. The asynchronous state, consistent with the traditional theory, corresponds to a stable fixed point, which can be destabilized through a Hopf bifurcation, giving rise to a stable limit circle, corresponding to network firing rate oscillation. As such, the E-I balanced state can incorporate network oscillation with different synchronous levels, which accounts for the coexistence of variability in both individual and population scales. In order to show that the coexistence of irregular spiking and collective critical avalanches is a widely existing phenomenon, we empirically verify this coexistence in up-states in the public experimental data of spontaneous spiking activities recorded in mouse somatosensory cortex in vitro (Ito et al., 2016). Scaling relations similar to the network model are also found to hold in these critical data sets. Our own analysis further indicates the universality of the observed phenomena in model networks. The theory proposed here explains how collective neural activities coexist with irregular neuron spiking and reveals a possible origin of criticality in neural systems. Compared with previous study (di Santo et al., 2018; Dalla Porta and Copelli, 2019) showing that critical avalanches exist around synchronous transition, our model incorporating synaptic kinetics, E-I balance, and irregular spiking of individual neurons provides a more biological plausible explanation about the nature of this transition. The theory also serves as a novel tool to study the dynamics of IF networks with biologically realistic synaptic filtering kinetics, and thus has useful application in large-scale modeling of brain networks.

**MATERIALS AND METHODS**

**Spiking Neuronal Network**

We study a leaky IF spiking neuronal circuit. Neurons are coupled by a random network with density \( p \) and size \( N_E + N_I \), which consists of \( N_E \) excitatory (E) neurons and \( N_I \) inhibitory (I) neurons. Thus, each neuron in the network has on average \( n_E = pN_E \) E neighbors and \( n_I = pN_I \) I neighbors. Each neuron also receives \( n_o \) excitatory inputs modelled by independent Poisson processes with frequency \( Q_o \), mimicking external inputs for the circuit under consideration. We set \( p = 0.2, N_E : N_I = 4 : 1, n_o = n_E \) and the network size is \( N = 10^4 \), unless otherwise specified. The sub-threshold membrane potential of neuron \( i \) at time \( t \), denoted as \( V_i(t) \), is governed by

\[
\frac{dV_i}{dt} = f_a(V_i) + J_{ao} \sum_{j \in \mathcal{N}_E} F_{i}^{E} \delta_j(t - \tau_E^j) + J_{AI} \sum_{j \in \mathcal{N}_I} F_{i}^{E} \delta_j(t - \tau_I^j) + J_{Al} \sum_{j \in \mathcal{N}_I} F_{i}^{E} \delta_j(t - \tau_I^j),
\]

(1)

where, \( V_i(t) \) is the membrane potential of neuron \( i \) (belonging to type \( \alpha = E, I \)) at time \( t \). \( \partial_{\alpha}^{\mathcal{N}} \) represents the \( \alpha \) neighbors of neuron \( i \). The input sources for neurons include excitatory inputs from external neurons (population \( O \)), inputs from recurrent excitatory neurons (population \( E \)) and inputs from recurrent inhibitory neurons (population \( I \)). The first term of Equation (1) describes the leaky current \( f_a(V_i) = (V_{\text{rest}}^E - V_i) / \tau_a \), which has the effect to drive the membrane potential back to the leaky potentials, which are set to be \( V_{\text{rest}}^E = V_{\text{rest}}^I = -70 \text{mV} \). The membrane time constants are set as \( \tau_E = 20 \text{ ms}, \tau_I = 10 \text{ ms} \) for E and I neurons, respectively. The second to fourth terms of Equation (1) are the external, excitatory recurrent and inhibitory recurrent currents, respectively. Input currents are the summations of the filtered pulse trains. Here, \( \delta_j(t) = \sum_n \delta(t - t_n^j) \) denotes the spike train of the \( j \)-th neuron. The excitatory and inhibitory synapses have latency period (delay) \( \tau_E^j \) and \( \tau_I^j \) respectively. For the numerical results presented, we consider \( \tau_E^j = \tau_I^j = 0 \) for simplicity (i.e., no transmission delay), which is a reasonable approximation for local circuits.

The synaptic filter is modelled as a bi-exponential function, i.e., Equation (2).

\[
F_{\alpha}^{\mathcal{N}}(t) = \frac{1}{\tau_{d}^{\mathcal{N}} - \tau_{r}} \left[ \exp \left( -\frac{t}{\tau_{d}^{\mathcal{N}}} \right) - \exp \left( -\frac{t}{\tau_{r}^{\mathcal{N}}} \right) \right], \quad t \geq 0.
\]

In Equation (2), we set the synaptic rise time \( \tau_r = 0.5 \text{ ms} \) for both E and I neurons, while the synaptic decay times \( \tau_{d}^{E}, \tau_{d}^{I} \) depend on the type of presynaptic neuron. We set \( \tau_{d}^{E} = 2 \text{ ms} \) and let \( \tau_{d}^{I} \) change from 1 to 4.5 ms to study the effect of different E and I synaptic filtering time scales. Hence, in our study here, \( \tau_{d}^{E} \) serves as a control parameter to induce the dynamical transition. Biologically, the inhibition decay time \( \tau_{d}^{I} \) depends on the constitution of synaptic receptors (Salin and Prince, 1996; Zhou and Hablitz, 1998), and can also be changed by chemicals such as narcotics (Brown et al., 2010). We point out that suitable changes of other model parameters such as synaptic strength, network connection density, etc., may also induce similar dynamical transitions we are going to study below. The integration dynamics is as follows. When the membrane potential reaches the threshold \( V_{th} = -50 \text{ mV} \), a spike is emitted and the membrane potential is reset to \( V_{\text{reset}} = -60 \text{ mV} \). Then, synaptic integration is halted for 2 ms for E neurons and 1 ms
for I neurons, modelling the refractory periods in real neurons. Synaptic weights are set as $J_{EO} = 0.45 \text{ mV}, J_{IO} = 0.72 \text{ mV}, J_{EE} = 0.36 \text{ mV}, J_{IE} = 0.72 \text{ mV}, J_{II} = -0.81 \text{ mV}$ and $J_{IH} = -1.44 \text{ mV}$, which will satisfy the balanced condition. Network dynamics are simulated by a modified second-order Runge-Kutta scheme (Shelley and Tao, 2001) with a time step of $dt = 0.05 \text{ ms}$. For each parameter, the network is simulated for $16 \text{ s}$ with the first $1 \text{ s}$ discarded to avoid a transient effect. The statistical indexes are then computed by averaging the results of $15$ trials with randomly distributed initial membrane potentials. The dynamics we considered here are current based. The case of conductance-based dynamics where the input of a neuron depends on its membrane potential is further studied in Supplementary Material: Appendix I. Model Extensions.

Mean-Field Theory of Network Dynamics Derivation of the Macroscopic Field Equations

We will develop a semi-analytical mean-field theory to approximate the average dynamics of the network by noting the fact that 1) the network topology is homogeneous and 2) the number of neuron $N$ is large. The following mean-field derivation holds for large enough $N$ and the network size is explicitly included in the field equations (see Equations (11), (12) below).

We denote the average membrane potential of the network as $V_{\alpha} = \langle V_{i} \rangle_{\alpha} := \langle V_{i} \rangle_{\alpha}, \alpha = E, I$. The goal is to derive a set of field equations governing the temporal dynamics of $V_{\alpha}$.

First, by noting that the convolution $F^{\alpha} \ast s_j(t - \tau^{\alpha}_j)$ obeys the equation

$$\left( V^{\alpha}_{\frac{d}{dt}} + 1 \right) \left( V^{\alpha}_{\frac{d}{dt}} + 1 \right) [F^{\alpha} \ast s_j (t - \tau^{\alpha}_j)] = \sum_{n} \delta(t - t^{\alpha}_n - \tau^{\alpha}_n),$$

(3)

then we have

$$\left( V^{\alpha}_{\frac{d}{dt}} + 1 \right) \left( V^{\alpha}_{\frac{d}{dt}} + 1 \right) \left[ \sum_{j \in E} F^{\alpha} \ast s_j (t - \tau^{\alpha}_j) \right]_{i \in E, I} \approx \sum_{j \in E} \sum_{n} \delta(t - t^{\alpha}_n - \tau^{\alpha}_n),$$

(4)

Under mean-field approximation, each neuron is the same in terms of their neighbors, so that

$$\left[ \sum_{j \in E} \sum_{n} \delta(t - t^{\alpha}_n - \tau^{\alpha}_n) \right]_{i \in E, I} \approx n_{\alpha} Q_{\alpha} (t - \tau^{\alpha}_n),$$

(5)

where $n_{\alpha}$ is the average number of $\alpha$ neighbors of a neuron in the network and $Q_{\alpha}(t)$ is the mean firing rate of $\alpha$ type neurons, defined as

$$Q_{\alpha}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{t}^{t + \Delta t} \left[ \sum_{n} \delta(s - t^{\alpha}_n) \right] ds.$$

(6)

In the standard definition of firing rate (Dayan and Abbott, 2001) of a neuron, the average in Equation (6) is taken over different simulation trials. Since ergodicity and the network homogeneity, neurons within a population should have the same firing rate and it can be computed through Equation (6) (i.e., population averages can be thought as sample averages) when the network is large enough in the stationary state. Formally, for measuring the firing rate from data, the time interval $\Delta t$ in Equation (6) has to be finite. Here we choose $\Delta t = 1 \text{ ms}$ (shorter than the refractory period) so that a neuron can at most have one spike between $t$ and $t + \Delta t$. Then, by the definition of $\delta$ function, $Q_{\alpha}(t)$ represents the proportion of $\alpha$ type neurons that spike between $t$ and $t + \Delta t$ as well as the mean firing rate of $\alpha$ type neurons at time $t$ with unit per $ms$.

In previous analysis framework of IF neurons through continuous stochastic processes theory (Burkitt, 2006), the membrane potential $V_i$ of neuron $i$ cannot cross the spiking threshold ($V_i$ is restricted to $(-\infty, V_{th})$ with $V_{th}$ being an absorbing boundary). This is a theoretical artefact, contrary to the true neurophysiology. Furthermore, in numerical integration, the resetting is achieved by finding those neurons whose membrane potential increases over the spiking threshold in each numerical step (Shelley and Tao, 2001). This inspires us to naturally consider that a neuron $j$ should have a spike at time $t$ if $V_j(t) > V_{th}$. Formally, we can consider $V_j(t)$ as the membrane potential of neuron $j$ at time $t$ before the resetting rule in each numerical step, then

$$Q_{\alpha}(t) = \langle H(V_j - V_{th}) \rangle_{j \in \alpha},$$

(7)

where $H$ is the Heaviside function $H(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$. Equation (7) explicitly builds the link that the population firing rate is the proportion of the neurons whose membrane potential is above the spiking threshold. As a preliminary approximation, we assume $V_j(t)$ obey a Gaussian distribution $P_{\alpha}(V)$ with time-dependent mean $V_{\alpha}(t)$ and time-independent variance $\sigma_{\alpha}^2$. We will verify that this assumption is plausible in the network and dynamic regimes we studied, referring to Figure 1D later. Then,

$$Q_{\alpha}(t) = \langle H(V_j - V_{th}) \rangle_{j \in \alpha} = \int_{V_{th}}^{\infty} P_{\alpha}(V) dV = \frac{1}{2} - \frac{1}{2} \text{erf} \left( \frac{V_{th} - V_{\alpha}}{\sqrt{2} \sigma_{\alpha}} \right),$$

(8)

where $\text{erf}$ is the error function $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt$. Although there is no elementary expression for the error function, it can be approximated by elementary functions. For example, a good approximation that can keep the first and second moments is $\text{erf}(x) \approx \tan(\frac{\pi x}{\sqrt{2}})$. Under this approximation, we have

$$Q_{\alpha}(t) = \frac{1}{2} - \frac{1}{2} \tan \left( \frac{\pi}{\sqrt{2}} \frac{V_{th} - V_{\alpha}}{\sigma_{\alpha}} \right) = \frac{1}{1 + \exp \left( \frac{V_{\alpha} - V_{th}}{\sigma_{\alpha}} \frac{\pi}{\sqrt{2}} \right)}.$$

(9)

Here, the standard deviation of the voltage, $\sigma_{\alpha}$, acts as an effective parameter to construct the voltage-dependent temporal firing rate. Note that this approximation scheme basing only on the first-order statistics neglects several factors that affect the accurate firing rate, including higher order statistics, noise correlation and refractory time. Thus, it does not have an analytical form and should be estimated numerically, from the steady-state mean voltage $V_{\alpha}^{ss}$ and mean firing rate $Q_{\alpha}^{ss}$ at the asynchronous state and
Equation (9) that
\[
\sigma_a = \frac{V_{th} - V_{ss}^a}{\ln((Q_a^o)^{-1} - 1)} \frac{\pi}{\sqrt{3}}.
\]

The sigmoid transfer function Equation (9) is the intrinsic non-linear property that induces oscillation transition in the field model. Note that neural field models of Wilson-Cowan type (Wilson and Cowan, 1972) would also contain a presumed sigmoid transfer function. Field models of this type can also qualitatively reproduce some dynamic features of E-I neuronal networks. Here we explicitly construct the sigmoid function from the microscopic spiking network. Thus, the quality of the scheme depends on suitable choices of effective parameters \(\sigma_a\) and once \(\sigma_a\) are chosen as suitable values, our field equations can predict the dynamics of the E-I network quantitatively (see Supplementary Material: Appendix 2. Sensitivity of the Critical Points on the Effective Parameters, where we study the sensitivity of the oscillation transition with respect to the effective parameters \(\sigma_a\)), while traditional neural field models (such as Wilson-Cowan model) are not quantitatively related to microscopic spiking networks in general.

Denote \(\Phi_a (t) = \left(\sum_{j \in I} F^a \ast s_j (t - \tau_a^j)\right)_{E,I}\) as the averaged synaptic time course of \(\alpha\) inputs received by a neuron and from Equations (4), (5), (9) it will obey
\[
\left(\tau_a^d \frac{d}{dt} + 1\right)\left(\tau_r \frac{d}{dt} + 1\right) \Phi_a = \left[\frac{n_a}{1 + \exp\left(\frac{V_{th} - V_{ss}^a (t - \tau_c^a)}{\sigma_a \sqrt{\pi}}\right)}\right],
\]
\[
\alpha = E,I.
\]

Note that each neuron receives \(n_a\) independent Poisson spike trains externally with rate \(Q_a\). Thus, the input of each neuron has a variance \(n_a Q_a\). If we do not consider its filtering effect since the fast decay time of excitatory synapse, by diffusion approximation we know the external input \(\sum_{j \in I} F^a \ast s_j (t - \tau_a^j)\) of each neuron can be approximated by \(n_a Q_a + \sqrt{n_a Q_a \xi_a (t)}\), where \(\xi_a (t)\) is a standard Gaussian white noise (GWN) with zero mean and unit variance. Since \(\{\xi_a (t)\}\) are independent GWNs, \(\sqrt{n_a Q_a \xi_a (t)}\) can be equivalently approximated by \(\sqrt{n_a Q_a / N_a} \xi_a (t)\), where \(\xi_a (t)\) is another standard GWN. Thus, taking the average \(\langle \cdot \rangle_a\) of the original Equation (1) and note that in the leaky IF model, \(f_\alpha\)
is linear, we arrive at

\[
\frac{dV_a}{dt} = f_a(V_a) + J_{ao}n_oQ_o + J_{ae}E_nQ_E + J_{ai}n_iQ_l = 0, \quad \alpha = E, I.
\]  

(13)

where \(Q_E, Q_I\) are given by Equation (9). In our case, the synaptic strengths and external input rates are the major parameters determining the value of the fixed point, while synaptic decay time would affect its stability. This is because the value of steady-state does not depend on \(\tau_d^E, \tau_r\), which can also be seen from Equation (2) that the synaptic filter is normalized \((\int_{0}^{\infty} F^S * \delta(t) = 1\) independent of \(\tau_d^E, \tau_r\)) so that synaptic rise and decay times would not affect the time-averaged firing rate. Thus, the scheme here cannot capture the nontrivial effect of synaptic filtering on affecting the firing rate [e.g., see the formula given by Fourcaud and Brunel (2002)]. Note that in general settings of balanced networks with dense and strong coupling (Renart et al., 2010), the quantities \(n_o, n_E, n_I\) are of order O(N) but the synaptic weights are of order O(N\(^{-1/2}\)). When N is large enough, the first term in Equation (13) can be neglected and it reduces to

\[
J_{ao}n_oQ_o + J_{ae}E_nQ_E + J_{ai}n_iQ_l = 0, \quad \alpha = E, I,
\]  

(14)

which is a set of linear equations to solve the steady firing rate \(Q_E, Q_I\). To guarantee a unique positive solution in this case, the sequence \((\frac{t^E_0}{N}, \frac{t^E}{N}, \frac{t^I}{N})\) should be in ascending or descending order, which is the so-called balanced condition in the traditional theory (Van Vreeswijk and Sompolinsky, 1996, 1998; Renart et al., 2010). Thus, the theory here is a generalization of the traditional theory of balanced network.

In the traditional theory of asynchronous dynamics, the E-I balanced state can be considered as the existence of a stable fixed point of Equations (11), (12). Now we can consider how the stability of this fixed point can be changed with the aid of this dynamic form. For the case without synaptic delay (i.e., \(\tau_d^E = \tau_d^I = 0\)), the field equations are ordinary differential equations. By taking \(X = (V_E, V_I, \Phi_E, \Phi_I, \frac{d\Phi_E}{dt}, \frac{d\Phi_I}{dt})^T\), the field model Equations (11), (12) can be written in the first-order form \(\frac{dX}{dt} = \mathbf{f}(X)\) without considering the noise. The Jacobian matrix at the steady state is

\[
\mathbf{J} = \begin{pmatrix}
-\frac{1}{\tau_E} & 0 & J_{EE} & 0 & J_{EI} & 0 \\
0 & -\frac{1}{\tau_I} & J_{IE} & 0 & J_{II} & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & n_E \frac{Q^I_E(V_E)}{\tau_d^E \tau_r} & 0 & -\frac{1}{\tau_d^E \tau_r} - \frac{1}{\tau_d^E} - \frac{1}{\tau_r} & 0 & 0 \\
0 & n_I \frac{Q^I_I(V_I)}{\tau_d^I \tau_r} & 0 & -\frac{1}{\tau_d^I \tau_r} & -\frac{1}{\tau_d^I} & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]  

(15)

with \(Q^I_a(V_a) = \frac{\pi \exp((V_a - V_o)\pi / (\sqrt{3} \sigma_o))}{\sqrt{3} \pi \sigma_o \left(1 + \exp\left((V_a - V_o)\pi / (\sqrt{3} \sigma_o)\right)\right)}\) estimated at the steady-state value of \(V_a\) given by Equation (13). The eigenvalues of \(\mathbf{J}\) can determine the stability of the steady state. Note that many models use single exponential function as the synaptic filter, i.e., \(\tau_r = 0\), and in this case the dynamic form becomes 4-dimensional with \(X = (V_E, V_I, \Phi_E, \Phi_I)^T\). For models without considering synaptic filtering effect (that is, the case of instantaneous synapse where \(\tau_r = \tau_d^E = 0\), the dynamic form becomes 2-dimensional, which can be considered as the dynamic form of traditional theory without synaptic kinetics. In these cases, the stability analysis can be performed in a similar way. In the presence of synaptic transmission delay, the field equations would become delay differential equations. In this case, stability analysis would in general become more difficult.

### Statistical Analysis of Model and Experiment Data

#### Spike Count Series

For statistically analyzing neural dynamics, the neuron spike train series have to be constructed from the model simulation data or the experimental data as follows. The time axis is first divided into consecutive time windows with sizes \(\Delta t\) ms. The number of spikes of neuron \(i\) is then counted in each window to obtain a discrete sequence \(N_i(t)\), which is designated as the spike
count series of neuron $i$ with time windows $\Delta t$. Alternatively, the number of spikes of the whole neuron population can be counted in each window. This constructs the population spike count series $N_{\text{E}}(t)$ for the $E$ and $I$ populations, respectively. Furthermore, $q_{\text{E}}(t) = \frac{N_{\text{E}}(t)}{n_{\Delta t}}$ is the population averaged firing rate series and the power spectrum of it can indicate the collective oscillation property. For computing certain quantities, such as the correlation between neurons, the spike count series is filtered by a square kernel $K_T(t) = \begin{cases} \frac{1}{T}, & t \in [-T, 0] \\ 0, & \text{other} \end{cases}$ with length $T$ and the ensuing filtered spike train is defined as $\hat{N}_i(t) = K_T * N_i = \frac{1}{T} \sum_{t=-T}^{T} N_i(t - s)$.

**Quantifying Spiking Irregularity and Firing Rate Variability**

The spiking time irregularity of a neuron can be quantified using the CV (coefficient of variance), which is defined as the standard deviation of the neuron ISIs (inter-spike intervals) over its mean. Totally regular activities have CV values of 0, while Poisson processes have CV values of 1. A higher CV value indicates larger irregularity. We measure the firing rate variability along the time on the individual neuron level and the population level by the relevant Fano Factor (FF). The FF of neuron $i$ is defined as $FF_i = \text{var}(N_i(t)) / \langle N_i(t) \rangle$. The average and the variance here are taken across the spike count series. Similarly, population FF is defined as $FF_p = \text{var}(N_p(t)) / \langle N_p(t) \rangle$. Note that FF depends on the time window size $\Delta t$ to construct the spike count series (we will use $\Delta t = 50 \text{ ms}$). The population FF can also be computed using the macroscopic field equations, as the field equations predicts the population firing rate, which can be transferred to population spike counts (by multiplying the number of neurons in the population).

**Quantifying Network Synchrony**

The synchrony of the network can be characterized from two aspects: the cross-correlation of spiking times and the coherence of the membrane potential (Golomb, 2007). The former quantifies the coherence of threshold events whereas the latter quantifies the coherence of the subthreshold dynamics.

We employ the commonly used Pearson correlation coefficient (PCC) to quantify the synchrony of the spiking time. The spike count series of neuron $i$ is first constructed with time window $\Delta t = 1 \text{ ms}$ and then filtered by a square kernel with length $T = 50 \text{ ms}$. The PCC between neuron $i$ and $j$ is defined as $\rho_{ij} = \frac{\text{cov}(n_i(t), n_j(t))}{\sqrt{\text{var}(n_i(t)) \cdot \text{var}(n_j(t))}}$. The details such as filtering in calculating PCC would affect its exact value (Cohen and Kohn, 2011), but not the qualitative change. The index $\{\rho_{ij}\}_{i,j \in p}$ PCC averaged over all excitatory neuron pairs, is used to quantify the network synchrony degree of threshold events.

The voltage series of each neuron is constructed for each millisecond $V_i(k) := V_i(t) \mid_{t=k \text{ ms}}$. Voltage coherence is defined as $\sqrt{\sigma^2 / \langle n^2 \rangle_{i,a}}$, where $\sigma^2 = \langle V^2_i \rangle - \langle V_i \rangle^2$ is the variance of the mean voltage $V_a = \langle V_i \rangle_{i \in a}$ and $\sigma^2 = \langle V^2_i \rangle - \langle V_i \rangle^2$ is the variance of the voltage of neuron $i$. The voltage coherence of the excitatory population is used to quantify the coherence of the subthreshold dynamics.

We further use the CV of the excitatory population firing rate series constructed with $\Delta t = 1 \text{ ms}$ to quantify temporal firing rate variability at short timescales, which is another way of indicating the network synchrony, as stronger synchrony indicates larger population firing rate variability at short timescales.

**Neuronal Avalanche Analysis**

We measure the neuronal avalanches of the excitatory neuron population from its population spike count series $N_{\text{E}}(t)$ constructed with window (bin) size $\Delta t$. An avalanche is defined as a sequence of consecutive non-empty bins, separated by empty bins (with no spiking inside). The size $S$ of an avalanche is defined as the total number of spikes within the period and the duration $T$ is defined as the number of time bins it contains. To compare different data sets in a unified standard, window size is chosen as the average ISI of the merged spiking train (constructed by merging the spike trains of all neurons), denoted as $T_{\text{m}}$. Thus, it depends on the mean firing rate of neurons in different data sets. This choice has been described as the ‘optimal’ window size to measure avalanches (Beggs and Plenz, 2003), as excessively small or large windows would lead to systematic bias. A further advantage of this choice is that the measured size and duration would approximately lie in the same scale ranges for different data sets, allowing better comparison. The effect of different choice of time windows for measuring avalanches is further considered in Supplementary Figure 3 and the results are robust for using time windows around $T_{\text{m}}$. Note that the definition of avalanche here is the time-binning (non-causal) avalanche, which corresponds to the case of experimental measurement (Beggs and Plenz, 2003; Bellay et al., 2015), but is different from the causal avalanche (Williams-Garcia et al., 2017) studied by many physical models rooted in critical branching processes.

To empirically judge the criticality of avalanche systems is a technically difficult issue. Inspired by the statistical property that avalanche size should obey a power-law distribution for scale-invariance, we adopt a two-step paradigm as follows.

First, we use a simple index to first roughly compare under which parameter the model is closer to criticality. The avalanche size frequency distribution histogram $(S, P(S))$, is first obtained with 80 plotting-bins from minimum to maximum size with points where $P(S) = 0$. The least squares method is then used to find the best-fit-line in doubly logarithmic coordinates, such that $\sum_i \left[ \log P(S) - (b_0 + b_1 \log S) \right]^2$ is minimised. After obtaining the best-fit coefficients $(b_0, b_1)$, the fitted frequency distribution values were estimated as $P_{\text{fit}}(S) = 10^{b_0 + b_1 \log S}$. Finally, the normalised distance, defined as $D = \sum_i S P(S) - P_{\text{fit}}(S) / \sum_i S P(S)$, the average size difference per avalanche between the actual and fitting frequency distribution normalized by the mean avalanche size, is used to measure the distance to the best fitting power-law distribution. The approach here is essentially the same as the widely used $\kappa$ index to assess the proximity to criticality in experiments (Shew et al., 2009). However, the use of $\kappa$ index requires a pre-definition of the slope of power-law (since it requires a form of the distribution...
function), which is not clear from data and was typically assumed to be $-1.5$ in some literature. Instead, the index we use is more convenient in that it does not require this pre-defined slope. This distance acts as an index for a rough comparison about which set is closer to criticality. However, it does not provide sufficient statistical evidence to claim criticality even when $D$ is very small.

Second, we use standard statistical methods to examine the critical properties when $D$ is relatively small. The maximum likelihood estimation (MLE) method by NCC toolbox (Marshall et al., 2016) is used to estimate the critical exponents. The toolbox provides a doubly truncated algorithm based on the MLE to find the range that passed the truncation based KS statistics test (Marshall et al., 2016). This truncation scheme can avoid the noise (in the small avalanche size range) and finite size effect (in the large avalanche size range) interruption in estimating the critical exponents. We find the largest truncated range that can pass the KS-based test with a $p$ value larger than 0.1 and broader than one-third of the whole range on the logarithm scale. This means that the data can produce a KS-statistics value that is less than the values generated by at least 10% of the power-law models in the truncated range. The estimated slopes within the truncated ranges in the avalanche size and duration distributions define the critical exponents $P(S) \sim S^{-\tau}$ and $P(T) \sim T^{-\alpha}$. A third exponent is defined as $(S(T)) \sim T^{1/\nu z}$, where $(S(T))$ is the average size of avalanches with the same duration $T$. This exponent is fitted using a weighted least squares method (Marshall et al., 2016) to those avalanches that fall into the truncated duration range for estimating $\alpha$. Finally, we examine the scaling relation (Equation (16) below) between these three exponents, which is a hallmark for criticality from the theory of statistical physics.

**Experimental Data Analysis**

We used the public experimental data measuring the neuronal spiking activity in mouse somatosensory cortex cultures in vitro (Ito et al., 2016). A total of 25 data sets were used and the length of each record is 1 h, with the exception of Set 19, which was 48 min long. The recordings were performed on organo-typical slice cultures after 2 to 4 weeks. The recordings were performed on organo-typical slice cultures after 2 to 4 weeks in vitro, without stimulation (Ito et al., 2014). Spiking times were sorted with a PCA-based algorithm (Litke et al., 2004) to locate the signals recorded with a large and dense multi-electrode array of 512 electrodes.

Under in vitro conditions, spiking in the culture clearly shows up-down state transition. Active spiking periods (up-state) and silent periods (down-state) alternated slowly with a frequency of circa 0.1 Hz. We focus on analyzing the up-state defined as follows. For each data set, the population firing rate series was first constructed with a time window of $\Delta t = 10$ ms and then filtered by a square kernel with length $T = 100$ ms. Then, the up-state is defined as the time periods that the population firing rates are higher than 30% of the maximum rate of the given dataset and last longer than 1 s. We further examine the power spectral density of the population firing rate series (before further filtering) in the up-state, the distribution of firing rates of the neurons, CV of ISIs and the avalanches in the up-states. The time bin for measuring the avalanche was the mean ISI, averaged through the up-state, of the merged spike train. To test critical properties, we first determine whether the size and duration distribution is close to a power-law and then estimate its critical exponents. As in our analysis of modelling data, the doubly truncated and statistical test algorithm from the NCC toolbox (Marshall et al., 2016) is used. We accept the power-law distribution of a data set if the following two conditions can be jointly satisfied: 1) the truncated range has to be broader than one-third of the whole range in the logarithm scale and; 2) the data in the truncated range can pass the KS-based test with $p$-value larger than 0.1. For the cases deemed to be power-law distribution, the truncated ranges, estimated exponents and $p$-values in the KS-based test are shown. For the data sets that have power-law avalanche distributions in both size and duration, we further compare the critical exponents $1/\nu z$ and the value $2/\nu z$ to see whether scaling relation (Equation (16) below) holds. To test whether power-law distribution is the intrinsic structure of the data, we also analyze the surrogate data constructed by randomly shuffling the ISIs of neurons in each up-state period. This random shuffling can destroy the intrinsic temporal correlation structures of the avalanches.

**RESULTS**

**Network Synchrony Arises From Loose Temporal Balance**

To begin, we first intuitively illustrate the non-trivial effect of synaptic filtering (Figure 1A) in shaping the network dynamics. For fast inhibition decay time, the network spiking dynamics is asynchronous (Renart et al., 2010), as shown in Figure 1B. Such a ‘strict balance’ is the case of traditional E-I balance theory (Van Vreeswijk and Sompolinsky, 1996, 1998) with inhibition domination, where network inhibitory feedback can cancel the excitatory current spill on a fast time scale (Figure 1H).

This scenario of strict balance breaks down when inhibition becomes slow, which induces an effective delay in the inhibitory cancellation (Figure 1I). Such a delay before the cancellation would induce a window during which excitatory current spills over the network, resulting in the collective spiking shown in Figure 1C. The strength of the collective activity depends on the length of this excitation-dominant window. As inhibition becomes dominant again after a delay, a temporal quiescent episode ensues after the collective spiking. Thus, slow synaptic inhibition induces network oscillation, as was also shown in previous studies (Brunel and Wang, 2003; Yang et al., 2017; Huang et al., 2019). In this ‘loose balance’ scenario, the network maintains balance on a slower time scale. In this case, as excitation and inhibition dominate alternately, the strict balance is temporally broken on a fast time scale. Note that in the loose balance state, the delay in the inhibitory cancellation is merely a few milliseconds (Figure 1I), consistent with the inhibition tracking delays observed in experiments (Okun and Lampl, 2008).

Alternatively, these different collective behaviors can be seen from the network population firing rate dynamics. The network firing rate in the strict balance, asynchronous state is almost steady (Figure 1E), whereas the loose balance induces firing
rate oscillation (Figure 1F), with fast firing rate variability at the population level. The emergence of oscillation implies a regularity of the population dynamics of the network. However, it is important to stress that the spiking of the neurons is still highly irregular. There are at least two reasons for this irregularity. First, the population oscillation is quasi-periodic rather than periodic due to the stochastic nature of the network dynamics. Second, and more importantly, each collective burst in the network is randomly participated in by a small portion (∼10% on average) of neurons in the network. These neurons spike with a firing rate much smaller than the network oscillation. This can also be understood from the fact that the firing rate is still low at the peaks of spiking activity corresponding to spiking synchrony (Figure 1D).

As such, the dynamical phenomenon here with loose balance was also referred to as ‘sparse synchrony’ (Brunel and Hakim, 2008). The histogram counts of the coefficient of variance (CV) of the inter-spike intervals (ISIs) of neurons (see Materials and Methods) in the network are shown in Figure 1G. In both asynchronous and synchronous states, the overall value of the CV is around 1, indicating that the spiking irregularity approximately resembles the Poisson process (Holt et al., 1996). As our mean-field theory assumes that the membrane potentials of the neurons in the network approximate Gaussian, we further analyze the distribution of the membrane potentials. As shown in Figure 1D, in both fast and slow inhibition cases the distribution is not totally Gaussian and instead right skewed, a feature of the finite threshold IF dynamic (Brunel, 2000; Keane and Gong, 2015). However, the skewness (defined as the third central moment divided by the cubic of the standard deviation) and kurtosis (defined as the fourth central moment divided by the quartic of the standard deviation) measuring the deviation from Gaussian are (−0.67, 3.12) and (−0.66, 3.3) in the fast and slow inhibition cases respectively, which are not far from (0, 3) for the case of Gaussian distribution. Thus, our theory that assumes a Gaussian distribution is still effective, as will be shown later.

**Mean-Field Theory Predicts the Synchronous Transition**

The network dynamic transition induced by a looser E-I balance can be characterized by the dynamics of population firing rate (Figures 1E,F). However, it is theoretically challenging to analyze the population dynamics of IF networks with synaptic kinetics. Here, we propose a novel mean-field approximation theory to derive the macroscopic dynamic equations of an IF network, i.e., Equations (11), (12), referring to a schematic diagram in Figure 2A. This technique for deriving the macroscopic field equations is highly generalizable. Extensions to the cases of time-varying inputs and conductance-based dynamics are presented in Supplementary Material: Appendix 1. Model Extensions.
For numerically estimated $\sigma_a$ (i.e., use Equation (10) to compute $\sigma_a$ under different external input strength $Q_o$), the mean firing rate of the network can be correctly estimated from the field Equations (11), (12) (Figure 2B blue). We further consider whether the field equations can predict the firing rate with fixed parameters $\sigma_a$. Following the derivation of Equation (12), if we simply assume $\frac{dV_i}{dt} \approx const - \frac{V_i}{\tau_o} + J_{ao} \sqrt{\alpha_o Q_o} \xi_i(t)$, then $V_i$ is Gaussian distributed with standard deviation $\sigma = \sqrt{\frac{1}{\tau_o} \alpha_o Q_o \tau_a}$. Figure 2B green shows the corresponding results of fixed $\sigma_o = \sqrt{\frac{1}{\tau_o} \alpha_o Q_o \tau_a}$ for $Q_o = 5$ Hz. It cannot predict the exact firing rate but can correctly predict the linear response to external input, a property of asynchronous balanced network (Van Vreeswijk and Sompolinsky, 1996).

The dynamic difference between the asynchronous state and synchronous state can already be predicted by the field equations in terms of the population firing rate dynamics, as shown in Figures 1E,F. The mechanism is explained by a Hopf bifurcation in the field equations through stability analysis (see Materials and Methods), as shown in Figure 2C. In the case of fast inhibition, the fixed point of the field model is generally a stable focus, whose Jacobian has complex eigenvalues with negative real parts. In this case, the network firing rate only fluctuates mildly due to noise perturbations. When $\tau_o$ increases, the fixed point will lose its stability through a supercritical Hopf bifurcation, as indicated by a pair of its dominant complex conjugate eigenvalues $\lambda = \alpha \pm i \omega$ crossing the imaginary axis. The Hopf bifurcation predicts that the stable fixed point will give way to a stable periodic solution, whose amplitude grows from zero. The frequency of the periodic motion can be estimated as $\omega = 2\pi$ in the linear order. This prediction is approximately equal to the numerically measured peak frequency of the network excitatory firing rate oscillation near the critical bifurcation point, as shown by the blue circles in Figure 2C. Previous studies have shown that E-I networks can undergo a transition to oscillation through perturbation analysis by assuming the form of the network steady-state firing rate solution (Brunel and Wang, 2003; Brunel and Hakim, 2008) and the phenomenon has been associated with a Hopf bifurcation of rate equations with effective transmission delay (Brunel and Hakim, 2008). However, there is no direct connection from the spiking networks to the heuristic rate equations in that theory. Here, our mean-field approach with the aid of macroscopic field equations derived from microscopic neuronal network straightforwardly reveals the Hopf bifurcation mechanism during this transition to sparse synchrony state. Note, however, that sparse synchrony states can also be achieved in purely inhibitory neuronal networks without synaptic kinetics (Brunel and Hakim, 1999).

As shown in Figure 2D, the network synchrony increases dramatically after $\tau_o^c$ crosses the Hopf bifurcation point. Figure 2E shows that the network temporal firing rate variability (see Materials and Methods) increases conspicuously during this transition, which is also qualitatively predicted by the field equations. The CV of ISIs averaged over the excitatory population is shown in Figure 2F. During the onset of collective oscillation, the averaged CV of ISIs first slightly decreases and then increases (as the strong bursting oscillation activity develops at around $\tau_o^c \approx 4$ ms, referring to Figure 3A bottom panel) while its overall value is around 1. The coexistence of irregular spiking and collective oscillation can be understood from the ratio between the mean firing rate of neurons and peak frequency of network oscillation. If each oscillation is participated by all the neurons, then the ratio should approach one. However, as can be seen from Figure 2F, the synchrony is sparse (Brunel and Hakim, 2008) in that this ratio is nearly 0.1 after oscillation onset, implying that each oscillation is randomly participated by only 10% of the neurons. Equivalently, each neuron only spikes sparsely and randomly participates in about 10% of the collective oscillations, giving rise to high variability of the ISI. As $\tau_o$ further increases to bursting onset, this ratio rapidly increases to approaching one, implying that almost all the neurons participate in each oscillation in the bursting state, where the inhibitory feedback is too slow such that the excitatory current can spill over the whole network in the excitatory dominant period.

In addition to spiking time variability, we further examine the fluctuations of firing rate in a short time scale (time window 50ms) by Fano Factor (FF) (Teich et al., 1997). As shown in Figure 2G, the change of FF of the individual neurons (see Materials and Methods) is similar to that of CV (Figure 2F), indicating that the individual firing rate fluctuation does not increase with the onset of collective synchrony. In contrast, the change of population FF (Figure 2G, and see Materials and Methods for details) is similar to the spiking correlation (Figure 2D) and temporal variability of firing rate (Figure 2E).

This trend was further qualitatively confirmed by the field model prediction, as shown in Figure 2G. Thus, it appeared that the observed collective oscillation is a network property rather than an individual neuron property. Hence, the network collective oscillation activity is compatible with individual neuron irregular spiking. The latter property is the prominent feature of E-I balance (Okun and Lampl, 2009).

In summary, strict balance with asynchronous network spiking is predicted by a stable focus whereas loose balance with sparsely synchronous network spiking is predicted by a stable limit circle in the field equations. Both strict and loose balance conditions support the irregular spiking of individual neurons.

**Scale-Free Neuronal Avalanches Near the Critical Transition Point**

For our mean-field theory of balanced networks, the Hopf bifurcation predicts that a global oscillation emerges in the field equations with oscillation amplitude growing continuously from zero. This picture is similar to a second-order phase transition in statistical physics when oscillation amplitude is taken as the order parameter. Thus, we measure the spiking avalanches (see Materials and Methods) to examine whether this kind of criticality can result in scale-invariant spiking behaviour, i.e., critical avalanches, as observed in experiments (Beggs and Plenz, 2003; Friedman et al., 2012; Bellay et al., 2015; Fontenele et al., 2019). Figure 3B illustrates the time course of an avalanche.

To more clearly compare the feature of avalanche dynamics, raster plots of the spiking time of excitatory neurons in three typical cases are shown in Figure 3A. These plots illustrate the asynchronous state (upper panel), the onset of synchrony and
oscillation (middle panel) and the developed collective oscillation (lower panel). Power spectrum density (PSD) analysis of network firing rate is shown in Figure 3D. The asynchronous state appears without collective oscillation (frequency peak) and the avalanche size and duration distributions are exponential-like, as shown by the blue curves in Figures 3E,F. It is thus subcritical. Alternatively, the asynchronous state can be considered as very noisy population oscillations that are participated by neurons very sparsely. The onset of synchrony is usually accompanied by a frequency peak of fast oscillation situated within the gamma bands, which is thought to have functional importance in various cognitive processes (Herrmann et al., 2004). This fast oscillation is temporally organized as scale-invariant avalanches, with power-law-like size and duration distributions, as shown by the green curves in Figures 3E,F. Figure 3D shows a peak at around 100Hz accompanied by its harmonics on top of a power-law decay $P(f) \sim f^{-\beta}$. Such a 'scale-free' behavior on PSD is observed in different local field potential data (He, 2014). By linear regression, we estimate that the decay exponent $\beta \approx 3.5$ in the critical state. Such decay in supercritical have a smaller exponent $\beta \approx 2$. Overall, these exponents lie in the typical range observed in local field potential data (such as EEG and EcoG) where $\beta$ ranges from 1 to 4 (He et al., 2010) and the value of beta is shown related to physiological or cognitive states (He et al., 2010).

To compare the critical properties of different dynamic states, we first examine the distance $D$ between the avalanche size distribution and its fitted power-law distribution (see Materials and Methods), an efficient way to roughly judge which dynamic state is closer to criticality, as power-law distribution is its most striking feature. Figure 3C shows that the distance $D$ is smallest when the network is poised near the Hopf bifurcation point predicted by the mean-field theory. Next, we confirmed the statistical significance of power-law distribution through truncated K-S test (see Materials and Methods) for dynamic states sufficiently close to the minimum of $D$. The existence of power-law distribution is only partial evidence of criticality, as other mechanisms could generate power-law distribution (Yadav et al., 2017). Finally, we further examined the scaling relation (Sethna et al., 2001) in the critical state (see Materials and Methods). The estimated slopes within the truncated ranges in the avalanche size and duration distributions define the critical exponents $P(S) \sim S^{-\tau}$ and $P(T) \sim T^{-\alpha}$. A third exponent is
defined as $\langle S \rangle (T) \sim T^{1/\alpha_x}$, where $\langle S \rangle (T)$ is the average size of avalanches with the same duration $T$. The third scaling feature can be found in both the subcritical and critical cases (Figure 3G), in accordance with previous experimental findings (Friedman et al., 2012). We find that the scaling relation (Sethna et al., 2001)

$$\frac{\alpha - 1}{\tau - 1} = \frac{1}{\sigma_{yz}}$$

(16)

approximately holds at the critical state. The exact value of critical exponents may depend on the details of the system. To demonstrate this, we slightly vary the input strength $Q_0$ from 4 to 8 Hz and the network still shows significant critical properties at the critical value $\tau_d = 3$ ms. As can be seen in Figure 3H, different exponents keeping the scaling relation Equation (16) can be found by varying the input strength $Q_0$. This phenomenon, together with the approximate linear relationship between exponents $\tau$ and $\alpha$ (inset of Figure 3H), is in accordance with the results measured in vivo in the primary visual cortexes (V1) of various animals (Fontenele et al., 2019) across a wide range of neural activity states, as well as our experimental data analysis later. The scaling relation expressed by Equation (16) provides additional evidence that the avalanches in the microscopic network occurring near the bifurcation point of the mean-field equations possess the properties of criticality. Interestingly, a recent theoretical study (Zierenberg et al., 2018) suggested that changing the external imposed input in vitro can tune the criticality of neural tissue (i.e., to make it approach or deviate from criticality) and it is latter confirmed in neuromorphic chips (Cramer et al., 2020). In line with this idea, our modeling study here predicts that around the critical states, slightly tuning the strength of external input can maintain the criticality of the network but modulate the critical exponents of the scale-free avalanches, which may be an interesting testable prediction for in vitro experiments.

Note that an avalanche can be understood as a temporal propagation of spiking activity in a network. These temporal propagations occur when the excitatory current temporally spill over the network and the avalanche size depends on the strength of the propagations, referring to Figures 1H, I. In the strict balanced state where excitation is cancelled by inhibition very fast, only small size avalanches can occur and the avalanche distribution is thus subcritical. On the contrary, in the loose balanced state with sparse network oscillations, large avalanches with typical scales can be induced by the temporal domination of excitation and the avalanche distribution can become supercritical. Only near the transition point where the macroscopic dynamic is also noise-driven, avalanches occur with all scales and the avalanche distribution thus can be scale-free. More specifically, on the macroscopic scale, the dynamical process of avalanche corresponds to a noise-induced excursion of the population firing rate. As our measurement of avalanches in the network by individual spiking times requires a fine time scale, and the information in this fine scale is averaged out in the field model, thus, information describing small avalanches vanishes in the field model, which only predicts the global firing rate dynamics. Although it is still difficult to directly link the microscopic avalanches dynamics to the macroscopic firing rate dynamics, the scale-invariance property of criticality inspires that the properties of burst activities (avalanche) in the macroscopic scale can shed light on the origin of power-law scaling in the microscopic network.

To consider the avalanche in the macroscopic scale, one should inspect the fluctuation behavior of a macroscopic signal $x(t)$ of the network, such as the mean firing rate, etc. An avalanche is a process that starts growing at $x(0) = x_{th} + \epsilon (\epsilon \to 0^+)$ and at the first time it goes back to $x(T) = x_{th}$ at time $T$. Here, $x_{th}$ is a threshold above which the avalanche is defined. The quantity $T$ turns out to be the first-passage time (FPT) (back to $x_{th}$) of this process and it defines the duration of an avalanche. The area $S = \int_0^T x(t)dt$ in this process defines the size of the avalanche. The scale-free behavior at the synchronous transition point may be understood as the general feature of dynamical systems near a bifurcation point when subjected to demographic noise $dx/dt \sim \sqrt{\xi(t)}$ (di Santo et al., 2017). This is because in the critical state of the E-I network, neurons are subjected to Poisson-like noise input with very weak correlation. Thus, according to a Gaussian approximation given by the central limit theorem, the overall fluctuation of population activity density scales with its square root as given by the central limit theorem (Benayoun et al., 2010; di Santo et al., 2018). Noise-driven random walker theory predicts that power-law distribution of the avalanche is a general feature of a dynamical system subjected to such kind of noise when near the Hopf bifurcation (actually, general bifurcations) point. Specifically, although the dynamic form of this macroscopic signal may not be explicit in the field equations, we can heuristically consider a situation that $X(t) = x - x_{th}$ obeys an intrinsic dynamic as the normal form of the amplitude dynamics of Hopf bifurcation but driven by noisy force modeled as GWN. The Langevin equation it obeys is

$$\frac{dX}{dt} = aX - X^3 + \eta (X, t).$$

(19)

The first part of Equation (19) is the normal form of the oscillation amplitude dynamics of a supercritical Hopf bifurcation (Marsden and McCracken, 2012), where periodic motion arises when $a$ increases across the bifurcation point $a = 0$. The second part $\eta (X, t)$ is the noisy driving force, where the fluctuation scales with square root of the activity. Thus, it has the form $\eta (X, t) = h + \sqrt{\xi(t)}$, where $h$ is the mean bias, including the effect from recurrent excitatory, recurrent inhibitory and external inputs. $\xi(t)$ is a standard GWN. The fact is that avalanche dynamics given by the first-passage process of Equation (19) can be mapped to the case of random walks in logarithmic potential (di Santo et al., 2017) by a scaling analysis (details in di Santo et al., 2017). Under this approach, the FPT distribution of the avalanche process can be solved by absorbing boundary approach in an analytical way, resulting in

$$P(T = t) = \frac{\sqrt{2\epsilon}}{\Gamma(\frac{1}{2} - h)} t^{\frac{1-h}{2}} \exp \left(-\frac{2\epsilon^2}{\tau} t \right).$$

(20)
FIGURE 4 | Coexistence of irregular spiking and critical avalanches in experimental data sets. (A) The raster plot of spiking time of data Set 1 in a period. Each point corresponds to a spike of a neuron. Indexes are sorted according to the total number of spikes during the whole recording time. The red/yellow segments in the top indicate the up/down states. (B) The distribution of the neurons firing rate in up and down states. (C) The power spectral density of the population firing rate in up and down states. In panels (B,C), the examples are shown by an up and a down period of data Set 1. (D) Boxplots of the CV distribution across neurons in different data sets. (E–G): critical avalanche properties of data Set 1. (E) Probability density distribution of the avalanche size. (F) Probability density distribution of the avalanche duration. (G) Mean avalanche size with respect to the given avalanche duration. The power-law distributions of size and duration are destroyed by surrogated data with shuffled ISI (denoted by blue curves). (H) Scaling relations between critical exponents of different data sets, similar to Figure 3H.

This explains the power-law distribution relation $P(T) \sim T^{-\alpha}$ with $\alpha = \frac{3}{2} - 2\theta$ and other exponents can be obtained by scaling argument (di Santo et al., 2017), although the relation between the avalanche critical exponents in the macroscopic scale and the microscopic scale needs further exploration. In all, irregular microscopic spiking leads to macroscopic fluctuation, which becomes the dominant effect that shapes the dynamic when near the macroscopic bifurcation point. The scale-free avalanche dynamics in the microscopic spiking network at the critical state may be understood as the scaling features of the first-passage dynamics near the Hopf bifurcation point in the macroscopic field model.

Coexistence of Irregular Spiking and Critical Avalanches in Experimental Data

In the following, we further analyze public experimental data measuring the in vitro neuronal spiking activity of mouse somatosensory cortex cultures (Ito et al., 2016) (25 data sets in total), to confirm whether the coexistence of irregular spiking and collective critical avalanches is a widely existent phenomenon.

For the in vitro experimental data, spiking in the culture clearly shows up-down state transition (Sanchez-Vives and McCormick, 2000; Luczak et al., 2007). Active spiking periods (up-state) and silent periods (down-state) alternate slowly with a frequency of circa 0.1 Hz, referring to Figure 4A. We focus on analyzing the up-state (see the definition in Materials and Methods) for the following reasons. First, neurons in up-states exhibit more spiking and stronger oscillation trends, which can be seen from the power spectral density (PSD) in Figure 4C. On the contrary, spikes were too few in the down-state, thus up-states are more likely to represent normal neural activities and closer to the dynamic regime we have studied in the model. Furthermore, the neurons recorded in data exhibit heterogeneity (Ito et al., 2014), i.e., broad distribution of individual firing rate, which is different from our modeling results with homogeneous random network topology. However, the spiking of neurons in up-state is more homogeneous, which can be seen from the evener distribution of firing rates of neurons (Figure 4B). Taken together, we expect that our modeling results may explain partial properties of the experimental data in up-states.
While most of the data sets display heavy tails in the distributions of avalanche size and duration (Supplementary Figures 4, 5), we find that nine of the 25 data sets exhibit both size and duration distributions that correspond to power-laws according to our standards (see Materials and Methods). Among them, seven sets contain critical exponents that approximately satisfy the scaling relation Equation (16) with errors < 0.1. Further details of the analysis results of the data sets are presented in Supplementary Table 1, which estimated the critical exponents of the data sets. These estimated critical exponents $\alpha$, $\tau$ and $\frac{1}{\tau}$ are also close to the ranges found in our model. The statistical results of these seven data sets are shown in Figures 4D,H. The boxplots of the distribution of CV of ISI across neurons in each set are shown in Figure 4D. Typical values of CV of neuron ISI are around 1~1.5, a range similar to the model prediction, indicating irregular spiking. Figures 4E–G illustrates the avalanche size and duration statistics of Set 1. Power-law distribution of avalanche size and duration can be observed. For surrogated data, where the inter-spike intervals of neurons were randomly shuffled, power-law distributions cannot be maintained, indicating that the critical properties are intrinsic in the data. As shown in Figure 4H, the scaling relationship between critical exponents (i.e., Equation (16)) approximately holds for these data sets. For these critical data sets, we also find a linear relationship between exponents $\tau$ and $\alpha$ (as seen in the inner panel of Figure 4H) whereby $\alpha = 1.36\tau - 0.37$, while the slope is slightly different from the value in our model with $\alpha = 1.21\tau - 0.154$ (inner panel, Figure 3H) and a previous study (Fontenele et al., 2019) with $\alpha = 1.28\tau - 0.28$. These results, obtained from experimental data, indicate the phenomena of the model that irregular neuron spiking can collectively organize as scale-invariant critical avalanches are also widespread in vitro, complementing the profound finding of this wide-spread phenomenon between different cortical states in vivo (Fontenele et al., 2019). Note that the model we study here does not have the property of slow transition between up and down states, which in general involves additional adaptation mechanisms (Jercog et al., 2017; Tartaglia and Brunel, 2017).

DISCUSSION

In summary, we have shown that E-I balanced IF networks with synaptic filtering kinetics can reconcile the coexistence of irregular spiking and collective critical avalanches near a synchronous transition point. The mechanism is unveiled by the macroscopic field equations derived by a novel mean-field theory which effectively capture the network dynamics. Finally, we further show that the phenomenon can be widely observed in the up-states of spontaneous spiking activities recorded in vitro in mouse somatosensory cortex cultures.

The E-I Balanced State Reconciles With Different Levels of Synchrony

Traditionally, E-I balance (Okun and Lampl, 2009) has been deemed to be the origin of neuron spiking irregularity. Classical mean-field theory (Van Vreeswijk and Sompolinsky, 1996, 1998; Renart et al., 2010) predicts that irregular spiking is an asynchronous state with low spiking time correlation. However, input synchrony (Stevens and Zador, 1998) has also been shown to contribute to the property of irregular spiking. This varied understanding renders a question on how to effectively characterize the dynamic of the E-I balance state.

Here, we study E-I networks that incorporates synaptic kinetics and the critical transition from strict to looser balance, where individual neurons continue to spike irregularly during these different dynamic states. Thus, individual irregular spiking is compatible with asynchronous (Renart et al., 2010) or sparse synchrony (Brunel and Hakim, 2008) state (i.e., synchronous inputs). Our theory shows that the balanced state can be a stable fixed point or a stable limit cycle on the macroscopic scale. The former induced by strict balance corresponds to the asynchronous state in accordance with the traditional theory, while the latter induced by loose balance corresponds to collective network oscillation, with critical dynamics during the transition between them. The critical dynamics where neurons are weakly synchronous also provides an explanation of how weak pairwise correlation can induce abundant collective behavior (Schneidman et al., 2006). Our study thus gives an effective characterization that the E-I balanced state can be a stable state of different characteristics (fixed point or limit cycle) in the macroscopic scale.

Possible Origin of Crackling Noise in Neural Systems Around Synchronous Transition Point

The criticality of neural systems has been long debated (Wilting and Priesemann, 2019a). Previous experimental and theoretical studies (Beggs and Plenz, 2003; Lombardi et al., 2012; Poil et al., 2012; Yang et al., 2012) have suggested that critical avalanches exist in the E-I balanced state, while many simplified models use unrealistic neural dynamics (such as spreading processes on networks) and the definition of E-I balance is usually ambiguous. Our study of IF network with realistic synaptic kinetics considers realistic asynchronous or synchronous spiking in E-I balanced state, and reveals that that critical avalanches exist near the synchronous transition point between these states, offering a more biologically plausible explanation of the origin of scale-free dynamics in neural systems.

Our theory is thus consistent with the understanding that criticality occurs at the edge of a synchronous transition (di Santo et al., 2018; Dalla Porta and Copelli, 2019; Fontenele et al., 2019) with intermediate levels of spiking variability (Fontenele et al., 2019), and that critical avalanches can temporally organize as collective oscillations (Gireesh and Plenz, 2008). It may appear that the co-existence of scale-free avalanches with scale-dependent network oscillations is contradictory. This can be understood by the accumulation-release processes of large avalanches where the system takes a typical period of time to accumulate before another large avalanche release, leading to a rhythm of large avalanches (Wang et al., 2016). Furthermore, the portion of large avalanches takes dominant effects on modulating the network oscillation, whereas the whole avalanches are still...
scale-free distributed. On the contrary, how these avalanches are modulated by individual oscillation of neurons may be another interesting problem for further exploration.

Some previous modeling studies also indicate that critical avalanches emerge near synchronous transitions. For example, (di Santo et al., 2018) studied spatially coupled units modeled by cubic normal form rate equations. Their physical model facilitates theoretical treatment but omits too many biological components (Dalla Porta and Copelli, 2019) studied a probability spiking model but without theoretical characterizing the transition dynamics. They focused on the analysis of population activity without considering the irregular spiking and E-I balanced input of individual units. Oscillation component can also be included in physical model of branching process (Pausch et al., 2020), while being away from the criticality in branching process means either explosion or being completely in salience, which is not biologically desirable. Thus, the biological reality of our model with synaptic filtering kinetics, E-I balance and irregular spiking time in individual neurons is well beyond these previous studies. Significantly, in our model, the synaptic transmission is not instantaneous due to the filtering effect, which renders difficulty in the distinction of the spatially causal relation (Martinello et al., 2017) between successive spiking. The time binning avalanches we measure here are temporally causal (as in experimental measurements) but not necessarily spatially causal in the network. Thus, our estimated critical exponents do not agree with the spatially causal avalanches produced by critical branching processes (Haldeman and Beggs, 2005) (i.e., directed percolation (DP) class), which predicts the classical results $\tau = 1.5, \alpha = 2, \frac{1}{\sigma_{\alpha z}} = 2$. Indeed, critical exponents different from the DP class have been found in previously reported experiments (Palva et al., 2013; Fontenele et al., 2019). Our own analysis of the experimental data (Figure 4H) also confirmed the variation of the exponents while maintaining the scaling relation Equation (16). Furthermore, the linear relationship between exponents $\tau$ and $\alpha$, also found in recent studies (Dalla Porta and Copelli, 2019; Fontenele et al., 2019), may depend on detailed properties of the underlying circuit and its dynamic origin remains to be further explored.

Another prediction of critical theory is that different avalanches can be collapsed into a single scale function (Sethna et al., 2001) at criticality. Such data collapse is a stricter criterion to probe criticality than the scaling relation Equation (16) we use here. It would be interesting to further explore whether there is a unique scale function that different avalanches can be collapsed into and how different oscillation components in the network can modulate its shape (Miller et al., 2019).

A Macroscopic Description of IF Neuronal Networks With Synaptic Kinetics

Traditional theory of E-I IF neuronal networks makes use of diffusion approximation (Amit and Brunel, 1997). For current-based dynamics without synaptic kinetics, Fokker Planck equation of membrane potential distribution can be constructed and firing rate can be derived by first-passage time theory (Amit and Brunel, 1997). Furthermore, the stability of the stationary membrane potential distribution can be analyzed through linear perturbation (Brunel, 2000), which can predict the transition from asynchronous spiking to sparse synchronous state through a Hopf bifurcation. On the contrary, our theory here analyzes stability of the dynamic of the mean membrane potential itself. Importantly, when synaptic kinetics is included, a complete analytical treatment does not exist. For example, the firing rate could only be derived asymptotically through a singular perturbation method by assuming fast decay synapses (Fourcaud and Brunel, 2002), whereas more ad hoc assumptions have to be further imposed for conductance-based dynamics (Brunel and Wang, 2001), and the oscillation properties of the network could only be analyzed by assuming a heuristic form of the network steady-state firing rate solution (Brunel and Wang, 2003). On the contrary, our semi-analytical theory can unifiendly treat networks with current-based or conductance-based (see Supplementary Material: Appendix 1. Model Extensions) dynamics with or without synaptic kinetics.

A more challenging but also more useful issue that has attracted much recent attention is to find macroscopic transient dynamic descriptions, e.g., rate equations, of the neuronal networks. Schaffer et al. (2013) derived the complex firing rate equations of IF networks through the eigenfunction expansion of the Fokker-Planck equation under diffusion approximation. Deriving rate equations of adaptive non-linear IF networks has also been studied (Augustin et al., 2017) under some effective approximation of the Fokker-Planck equation. Montbrió et al. (2015) derived the rate equations for quadratic IF networks using the Lorentzian ansatz. This approach has been generalized to including gap junctions (Laing, 2015) and synaptic filtering kinetics (Dumont and Gutkin, 2019). Schwalger et al. (Schwalger et al., 2017) developed a method to derive the stochastic rate equations of adaptive IF networks based on mean-field approximation of the renewal equation. In general, analytical theory works for specific conditions and thus is highly specific and with strong complexity. Furthermore, most proposed theories failed to capture the synchronous transitions induced by the synaptic filtering effect studied in our model.

Instead of being theoretically perfect, here we seek for simplicity and effectiveness. The key novel feature of our framework (see detailed derivations in Materials and Methods) is that the mean-field equations of the macroscopic dynamical variables can be closed by the voltage-dependent mean firing rate (i.e., Equation (9)). While many previous theoretical analyses of E-I network usually require the assumption of low input correlation (Van Vreeswijk and Sompolinsky, 1998; Brunel, 2000), the formula here is constructed by counting the number of neuron spiking in a small time window but not required to explicitly consider the correlation of neurons, which allows it to essentially capture the sub and supra threshold microscopic dynamics of a spiking network. Our novel scheme to derive the dynamic equations governing the first-order statistics of the neural network is not completely analytical, since the derivation neglects several factors that shape the network dynamics, including higher order statistics, noise correlation and refractory time. All these neglected factors have
been incorporated into the effective parameters, $\sigma_\alpha$, given by Equation (10). Nonetheless, this semi-analytical theory provides a simple yet useful method to analyze dynamic features directly related to the network firing rate, such as synchrony, criticality and response to dynamic stimuli. The semi-analytical nature of the theory results in the simple form of the macroscopic field equations, which facilitates further generalization. Previous work (Wong and Wang, 2006) proposed another semi-analytical theory for deriving rate equation from spiking neural networks. That method is a direct generalization of the classical Wilson-Cowan model (Wilson and Cowan, 1972) with presumed I-f curves (the relationship between input current and output rate). The derived field equations include the synaptic gating variables and thus can capture the synaptic filtering effect. Yet, unlike our approach here, that analysis cannot directly predict the dynamical properties of the spiking network (e.g., criticality, network oscillations) from the derived field equations.

**Outlook**

It is straightforward that the mean-field analysis introduced here can be used in the study of neural dynamics across a diverse range of topics. It can be easily generalized to include other factors by assuming different macroscopic variables in the field model. Possible extensions of the analysis include studying the effects of multiple neuronal populations and synaptic receptor types, cluster or spatially extended network connections, adaptive behaviors such as short-term plasticity, etc. Thus, our work allows further exploration of the mechanism that determines the role of synaptic kinetics in working memory retrieval (Mongillo et al., 2008), self-organized critical phenomena due to plasticity (Levina et al., 2007; Millman et al., 2010) and spatially causal avalanches or waves (Keane and Gong, 2015; Keane et al., 2018; Gu et al., 2019) arising from competition between Hopf and Turing instability (Huang et al., 2019). As a theory that links microscopic neuronal spiking and macroscopic collective activity that are consistent in several aspects with real neural dynamics with regard to E-I balance and neural criticality, our theory also establish a base to model large-scale brain connectomes (Haimovici et al., 2013; Chaudhuri et al., 2015; Wang et al., 2019), to study large-scale brain networks and information processing with realistic multiscale complex dynamics. In the experimental data analysis, we have confirmed the wide coexistence of irregular spiking and critical avalanches in the up-states of in vitro neuronal cultures, but our model could not reproduce some features in the experimental data sets: the uneven distribution of firing rates of neurons may require a heterogeneous network topology; the up-down states transition may require adaptation mechanisms in a slower time scale (Jercog et al., 2017; Tartaglia and Brunel, 2017). It would be interesting to further generalize our theory to models with such features. The potential applications of our theory listed here will receive the further exploration they deserve in due course.

**DATA AVAILABILITY STATEMENT**

Publicly available datasets were analyzed in this study. This data can be found here: http://crcns.org/data-sets/ssc/ssc-3.

**AUTHOR CONTRIBUTIONS**

JL and CZ conceived the study. TZ and CZ supervised the project. JL performed the theoretical analysis, numerical simulation, and data analysis. JL, TZ, and CZ wrote the manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

This work was supported by the Hong Kong Baptist University (HKBU) Strategic Development Fund, the Hong Kong Research Grant Council (GRF12200217), the HKBU Research Committee and Interdisciplinary Research Clusters Matching Scheme 2018/19 (RC-IRCMs/18-19/SCI01), the National Natural Science Foundation of China (Grants Nos. 11775314, 91530320, and 11975194), and Key-Area Research and Development Program of Guangzhou (Grants No. 202007030004). This research was conducted using the resources of the High-Performance Computing Cluster Centre at HKBU, which receives funding from the RGC and the HKBU.

**ACKNOWLEDGMENTS**

This manuscript has been released as a pre-print at arXiv.org (Liang et al., 2020). The authors would like to thank John M. Beggs for helpful discussion and pointing us to the public data sets (Ito et al., 2016). We also thank all reviewers for their comments and suggestions.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnsys.2020.580011/full#supplementary-material

**REFERENCES**

Abeles, M. (1991). Corticonics: Neural Circuits of the Cerebral Cortex. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511574566

Amit, D. J., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb. Cortex 7, 237–252. doi: 10.1093/cercor/7.3.237

Augustin, M., Ladenbauer, J., Baumann, F., and Obermayer, K. (2017). Low-dimensional spike rate models derived from networks of adaptive integrate-and-fire neurons: comparison and implementation. PLoS Comput. Biol. 13:e1005545. doi: 10.1371/journal.pcbi.1005545

Beggs, J. M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. J. Neurosci. 23, 11167–11177. doi: 10.1523/JNEUROSCI.23-35-11167.2003
organization of neuronal avalanches. *Phys. Rev. Lett.* 108:228703. doi: 10.1103/PhysRevLett.108.228703

Luczak, A., Barthó, P., Marguet, S. L., Buzsáki, G., and Harris, K. D. (2007). Sequential structure of neocortical spontaneous activity in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 104, 347–352. doi: 10.1073/pnas.060564104

Ma, Z., Turritanino, G. G., Wessel, R., and Hengen, K. B. (2019). Cortical circuit dynamics are homeostatically tuned to criticality in vivo. *Neuron* 104, 655–664. doi: 10.1016/j.neuron.2019.08.031

Marsden, J. E., and McCracken, M. (2012). *The Hopf Bifurcation and Its Applications*. Berlin: Springer Science & Business Media.

Marshall, N., Timme, N. M., Bennett, N., Ripp, M., Lautenhuber, E., and Beggs, J. M. (2016). Analysis of power laws, shape collapses, and neural complexity: new techniques and matlab support via the NCC toolbox. *Front. Physiol.* 7:250. doi: 10.3389/fphys.2016.00250

Martinetto, M., Fidzgal, J., Maritan, A., Di Santo, S., Plenz, D., and Muñoz, M. A. (2017). Neutral theory and scale-free neural dynamics. *Phys. Rev. X* 7:41071. doi: 10.1103/PhysRevX.7.041071

Miller, S. R., Yu, S., and Plenz, D. (2019). The scale-invariant, temporal profile of neuronal avalanches in relation to cortical γ-oscillations. *Sci. Rep.* 9, 1–14. doi: 10.1038/s41598-019-52326-y

Millman, D., Mihalas, S., Kirkwood, A., and Niebur, E. (2010). Self-organized criticality occurs in non-conservative neuronal networks during up states. *Nat. Phys.* 6:601. doi: 10.1038/nphys1757

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Sympatic theory of working memory. *Science* 319, 1543–1546. doi: 10.1126/science.1150769

Monnibrit, E., Pazó, D., and Rosin, A. (2015). Macroscopic description for networks of spiking neurons. *Phys. Rev. X* 5:21028. doi: 10.1103/PhysRevX.5.021028

Okun, M., and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat. Neurosci.* 11:535. doi: 10.1038/nn.2105

Okun, M., and Lampl, I. (2009). Balance of excitation and inhibition. *Scholarpedia* 4:7467. doi: 10.4249/scholarpedia.7467

Palva, J. M., Zhigalov, A., Hirvonen, J., Korhonen, O., Linkenkaer-Hansen, K., and Palva, S. (2013). Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3585–3590. doi: 10.1073/pnas.1216855110

Pausch, J., Garcia-Millan, R., and Pruessner, G. (2020). Time-dependent branching processes: a model of oscillating neuronal avalanches. *Sci. Rep.* 10, 1–17. doi: 10.1038/s41598-020-69705-5

Poli, S.-S., Hardstone, R., Mansvelder, H. D., and Linkenkaer-Hansen, K. (2012). Critical-state dynamics of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal networks. *J. Neurosci.* 32, 9817–9823. doi: 10.1523/JNEUROSCI.3990-11.2012

Renart, A., De La Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., et al. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590. doi: 10.1126/science.1179850

Rosenbaum, R., Smith, M. A., Kohn, A., Rubin, J. E., and Doiron, B. (2017). The asynchronous state in cortical circuits. *Sci. Rep.* 7:11016. doi: 10.1038/s41598-017-12033-y

Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T., and Kaplan, E. (1997). Fractal character of the neural spike train in the visual system of the cat. *J. Opt. Soc. Am. A* Opt. Image Sci. Vis. 14, 529–546. doi: 10.1364/JOSAA.14.000529

Touboul, J., and Destexhe, A. (2017). Power-law statistics and universal scaling in the absence of criticality. *Phys. Rev. E* 95:124134. doi: 10.1103/PhysRevE.95.012413

Van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726. doi: 10.1126/science.274.5293.1724

Van Vreeswijk, C., and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput.* 10, 1321–1371. doi: 10.1162/089976698300017214

Wang, R., Lin, P., Liu, M., Wu, Y., Zhou, T., and Zhou, C. (2019). Hierarchical connectome modes and critical state jointly maximize human brain functional diversity. *Phys. Rev. Lett.* 123:38301. doi: 10.1103/PhysRevLett.123.038301

Wang, S.-J., Hilgetag, C., and Zhou, C. (2011). Sustained activity in hierarchical modular neuronal networks: self-organized criticality and oscillations. *Front. Comput. Neurosci.* 5:30. doi: 10.3389/fncom.2011.00030

Wang, S.-J., Ouyang, G., Jiang, C., Zhang, M., Wang, Y. M., and Zhou, C. (2016). Stochastic oscillation in self-organized critical states of small systems: sensitive resting state in neural systems. *Phys. Rev. Lett.* 116:18101. doi: 10.1103/PhysRevLett.116.018101

Williams-Garcia, R. V., Beggs, J. M., and Ortiz, G. (2017). Unveiling causal activity of complex networks. *Europhys. Lett.* 119:18003. doi: 10.1209/0295-5075/119/18003

Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24. doi: 10.1016/S0006-3495(72)86068-5

Wilting, J., and Priesemann, V. (2019a). 25 years of criticality in neuroscience: established results, open controversies, novel concepts. *Curr. Opin. Neurobiol.* 58, 105–111. doi: 10.1016/j.conb.2019.08.002

Wilting, J., and Priesemann, V. (2019b). Between perfectly critical and fully irregular: a reverberating model captures and predicts cortical spike propagation. *Cereb. Cortex* 29, 2759–2770. doi: 10.1093/cercor/bhz049

Wong, K.-F., and Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328. doi: 10.1523/JNEUROSCI.3733-05.2006

Xue, M., Atallah, B. V., and Scanziani, M. (2014). Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature* 511, 596. doi: 10.1038/nature13321

Yadav, A. C., Ramaswamy, R., and Dhar, D. (2017). General mechanism for the 1/f noise. *Phys. Rev. E* 96:022215. doi: 10.1103/PhysRevE.96.022215

Yang, D.-P., Zhou, H.-J., and Zhou, C. (2017). Co-emergence of multi-scale cortical activities of irregular firing, oscillations and avalanches achieves cost-efficient
information capacity. *PLoS Comput. Biol.* 13:e1005384. doi: 10.1371/journal.pcbi.1005384

Yang, H., Shew, W. L., Roy, R., and Plenz, D. (2012). Maximal variability of phase synchrony in cortical networks with neuronal avalanches. *J. Neurosci.* 32, 1061–1072. doi: 10.1523/JNEUROSCI.2771-11.2012

Zhou, F.-M., and Hablitz, J. J. (1998). AMPA receptor-mediated EPSCs in rat neocortical layer II/III interneurons have rapid kinetics. *Brain Res.* 780, 166–169. doi: 10.1016/S0006-8993(97)01311-5

Zierenberg, J., Wilting, J., and Priesemann, V. (2018). Homeostatic plasticity and external input shape neural network dynamics. *Phys. Rev. X* 8:31018.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Liang, Zhou and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.