Topology-dependent coalescence controls scaling exponents in finite networks

Roxana Zeraati,[1,2,*] Victor Buendía,[1,2,*] Tatiana A. Engel,[3,4] and Anna Levina[1,2,5,†]

[1]*University of Tübingen, Tübingen 72076, Germany*
[2]*Max Planck Institute for Biological Cybernetics, Tübingen 72076, Germany*
[3]*Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey 08540, USA*
[4]*Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA*
[5]*Bernstein Center for Computational Neuroscience Tübingen, Tübingen 72076, Germany*

(Received 15 November 2022; accepted 26 February 2024; published 6 May 2024)

Studies of neural avalanches across different data modalities led to the prominent hypothesis that the brain operates near a critical point. The observed exponents often indicate the mean-field directed-percolation universality class, leading to the fully connected or random network models to study the avalanche dynamics. However, cortical networks have distinct nonrandom features and spatial organization that is known to affect critical exponents. Here we show that distinct empirical exponents arise in networks with different topology and depend on the network size. In particular, we find apparent scale-free behavior with mean-field exponents appearing as quasicritical dynamics in structured networks. This quasicritical dynamics cannot be easily discriminated from an actual critical point in small networks. We find that the local coalescence in activity dynamics can explain the distinct exponents. Therefore, both topology and system size should be considered when assessing criticality from empirical observables.

DOI: 10.1103/PhysRevResearch.6.023131

I. INTRODUCTION

Brain activity displays a plethora of different dynamical states, including bursts, oscillations, and irregular activity. In particular, neural activity exhibits spatiotemporal patterns compatible with the dynamics of a system close to a second-order phase transition, namely the critical point [1–5]. Operating at this regime has been linked to optimal information processing [2,6], efficient coding [7], maximal dynamic range and sensitivity to stimulus [8,9], longer timescales during attention [10], and better stimuli discrimination [11,12]. Neural network models demonstrated that short- and long-term synaptic plasticity can self-organize brain dynamics towards a critical point [13,14].

To assess whether the brain operates at criticality, the activity propagation between the neurons is often mapped to the branching process [1–5,8,15–17]. This mapping was motivated by the observation of outbursts of neural activity *in vitro* known as *neuronal avalanches* [1]. Each avalanche consists of periods of activity, separated by quiescence moments. The size and duration of neuronal avalanches in cortical cultured slices were found to follow power-law distributions with exponents $\tau = 1.5$ (for the size distribution) and $\alpha = 2$ (for the duration distribution), which are compatible with the avalanche exponents of the branching process at the critical point. The branching process dynamics is fully characterized by a branching parameter m [18], with $m = 1$ being the critical value, where each neuron, on average, activates one other neuron, creating a fluctuation-driven regime. From the statistical mechanics point of view, the critical transition at $m = 1$ belongs to the *mean-field directed percolation* (MF-DP) universality class, a nonequilibrium phase transition separating absorbing and active phases [19,20].

However, mapping the neural activity propagation to the branching process neglects the role of underlying network topology in shaping the dynamics. The branching process assumes the nonoverlapping spreading of the activity. In biological neural networks, in contrast, each neuron can be simultaneously excited by multiple sources. This phenomenon, known as *coalescence*, renders the independence assumption invalid and reduces the effective branching parameter of the system, since some active neurons cannot trigger spikes in already excited neighbors [17]. These effects are particularly severe in structured networks, such as in the primate cortex [21–26]. Additionally, the theory of critical phenomena predicts that structured, finite-dimensional network topology affects variables such as critical exponents [19,27].

To investigate the relationship between network topology and critical dynamics, we developed a finite-size branching network model with various connectivity structures, ranging from spatially arranged networks resembling the local connectivity structure of the cortex, to random or all-to-all connectivity. We show that the network topology controls the critical branching parameter, avalanche-size distributions, and their exponents.

*These authors contributed equally to this work.
†Corresponding author: anna.levina@uni-tuebingen.de

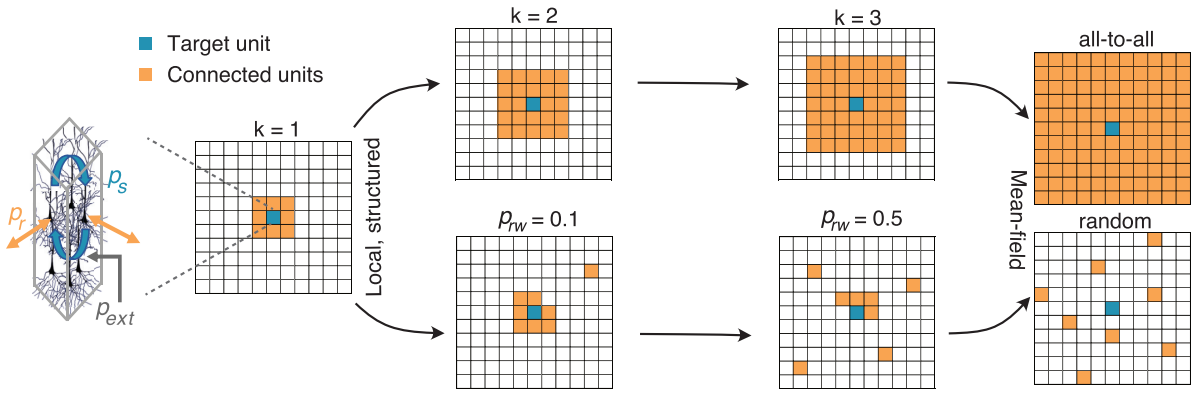

FIG. 1. Branching network model with different types of spatial connectivity. Each unit represents a cortical column (left) that can be excited by the self-excitation (probability p_s, blue arrows), neighboring units (probability p_r, orange arrows), and the external input (probability p_{ext}, gray arrow). k is the connectivity radius and p_{rw} is the rewiring probability. We consider networks with connectivity structures ranging from local (structured, left) to mean-field (random or all-to-all, right) generated via two pathways: by increasing the connectivity radius k (top) or rewiring local connections to random with the rewiring probability p_{rw} (bottom).

II. BRANCHING NETWORK MODEL WITH STRUCTURED CONNECTIVITY

The model consists of binary units arranged in a two-dimensional ($L\times L$) square grid with periodic boundary conditions. The connectivity is defined by two parameters: connectivity radius k and the probability of rewiring p_{rw}. First, each unit is connected to all units in its k-Moore neighborhood [$(2k+1)^2-1$ neighbors]. Then, with probability p_{rw}, each connection can be selected for rewiring and rewired to a uniformly chosen random location. We consider a network with $k=1$ and $p_{rw}=0$ and then systematically increase either the radius (Fig. 1, top) or the rewiring probability (Fig. 1, bottom). Increasing p_{rw} is equivalent to creating a two-dimensional (2D) small-world network as in the Watts-Strogatz procedure [28]. For $k=L/2$, a network with size $N=L^2$ will be all-to-all connected without any structure. For $p_{rw}=1$, we obtain a completely randomly connected network. Both limit cases correspond to the mean-field configuration.

Each unit i in the network transitions stochastically between an active state $s_i=1$ and an inactive state $s_i=0$, depending on the connectivity and external input:

$$p(s_i=0\to 1)=1-(1-p_{\text{ext}})(1-p_r)^{\sum_{j\in\Omega_i}s_j}, \quad (1)$$

$$p(s_i=1\to 0)=(1-p_{\text{ext}})(1-p_s)(1-p_r)^{\sum_{j\in\Omega_i}s_j}, \quad (2)$$

where p_r is the probability of being excited by the active neighbor, p_{ext} is the probability of receiving external input, Ω_i represents the set of neighbors of the ith unit, and p_s is the probability of maintaining the active state. This model is inspired by the activity and interactions in the cortex: the units represent cortical columns, and the active and inactive states correspond to transient high and low levels of activity found in the primate visual cortex [29,30]. The probability p_s accounts for vertical recurrent excitation between neurons within one column, and p_r represents horizontal recurrence between columns (for more details, see [10,31]). For simulations, we assume a timescale-separated regime with no external input ($p_{\text{ext}}=0$), and we take $p_s=0.5$ (unless stated otherwise). In the absence of external input, the transition matrix defining the dynamical equations can be simplified as

$$P=\begin{pmatrix} 1-F(x) & F(x) \\ (1-p_s)[1-F(x)] & p_s+(1-p_s)F(x)\end{pmatrix}, \quad (3)$$

where

$$F(x)\equiv p(s_i=0\to 1)=1-(1-p_r)^{\sum_{j\in\Omega_i}s_j}. \quad (4)$$

The model is a spatially structured version of a branching network (BN) [8,32]. The BN with random connectivity is completely described by local branching parameter $m=p_s+|\Omega_i|p_r$, summing all the outgoing connection probabilities of one node. On average, the number of active units A_t at time t when there was a single active unit at the previous time step is $\mathbb{E}[A_t|A_{t-1}=1]=m$. By taking $p_r=(m-p_s)/\langle|\Omega_i|\rangle$ and $p_s=0.5$ we reparametrize the model in terms of m. With a constant m, increasing connectivity k reduces the strength of individual connections p_r, while increasing rewiring probability p_{rw} does not affect p_r.

III. CRITICAL AND QUASICRITICAL DYNAMICS IN STRUCTURED NETWORKS

We first analyze the location of the critical transition depending on network topology using m as the control parameter. The exact location of the critical transition can be found by several methods (Appendix A). Here, we look for the critical branching parameter m_c that maximizes the variance of the activity $\chi(\rho(m))$. The simulations start with randomly activating a small percentage of units (15%; more details are in Appendix A).

In the mean-field system, the absorbing-active transition happens at $m_c=m_{\text{MF}}=1$ [20]. However, we find that the location of the critical transition depends on the network topology, in agreement with the theory of critical phenomena [27,33] (Fig. 2) and previous findings for small-world networks [34,35]. For the structured connectivity ($k=1$), the phase transition happens at a larger critical branching parameter ($m_c=1.109$). As we move towards the mean-field connectivity (either all-to-all or random), the critical

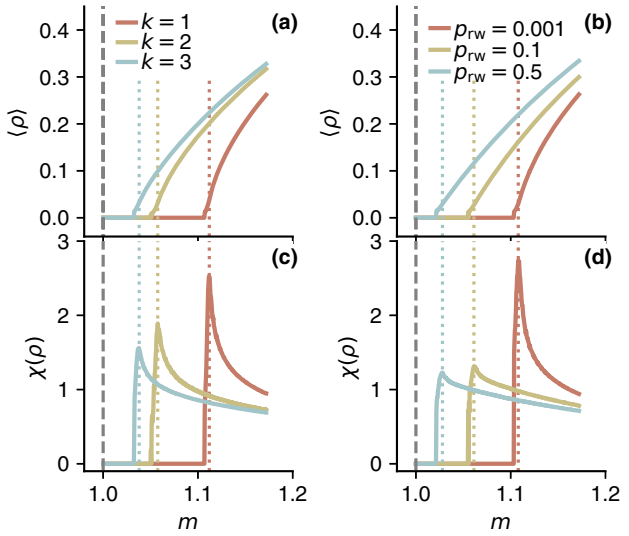

FIG. 2. Location of the critical transition depends on the network topology. Criticality occurs at the transition to the nonzero mean activity $\langle\rho\rangle$ (a), (b) and maximal variance $\chi(\rho)$ (c), (d), represented by vertical dotted lines. With increasing connectivity radius k (a), (c) or rewiring probability p_{rw} (b), (d), the critical branching parameter moves to the mean-field value (vertical gray dashed line). For simulations, $L = 128$. $p_{\mathrm{rw}} = 0$ for (a), (c), $k = 1$ for (b), (d).

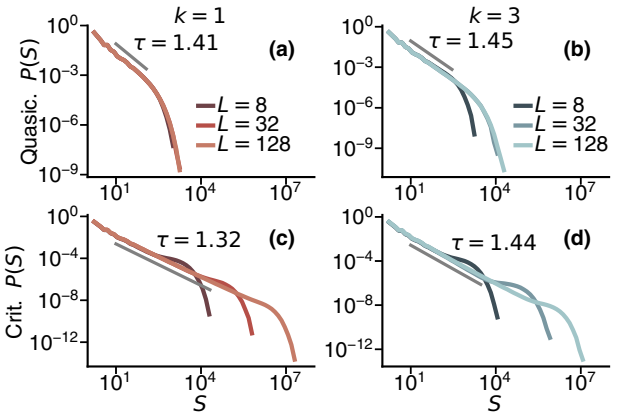

FIG. 3. Avalanche-size distributions differ in quasicritical and critical networks. System-size dependence at quasicriticality (a), (b) and criticality (c), (d), for $k = 1$ (a), (c) and $k = 3$ (b), (d). For critical systems, the cutoff location of the avalanche-size distribution shifts with system size (L^2) (c), (d). Quasicritical avalanches follow a power law up to a cutoff that is scaling with the system size for small systems (see also Fig. 10) but exhibit a characteristic scale for large systems (a), (b). Gray lines indicate the fitted power-law distribution with exponent τ for $L = 128$ (see Appendix D for fitting details; estimated exponents have an error margin of ± 0.01). m_c is defined as in Fig. 2.

branching parameter gradually converges to $m_{\mathrm{MF}} = 1$ (Fig. 2). In the structured networks, we refer to the dynamics at the mean-field branching parameter ($m_{\mathrm{MF}} = 1$) as quasicritical. In the absence of self-excitation ($p_s = 0$), the location of the critical point moves slightly, but our conclusions regarding critical and quasicritical dynamics stay the same (Appendix B).

The location of m_c is model-dependent. For example, in the two-dimensional contact process (CP) the critical point is located at $m_{\mathrm{CP}} \approx 1.6$ [20], while in our structured model ($k = 1$) criticality appears around $m_c \approx 1.1$ (Appendix C).

Next, we compare the avalanche-size distributions—often seen as the primary indicator for critical behavior in neural data—between the quasicritical (with m_{MF}) and critical (with m_c) networks for various network structures and sizes. An avalanche is a cascade of activity propagation in the network. It starts with an external input activating a single neuron in a quiescent network and ends when the activity dies out. At criticality, avalanche sizes and durations follow a power-law distribution with an exponential cutoff whose location scales with the network size. We compute the avalanche statistics in the timescale-separated regime ($p_{\mathrm{ext}} = 0$): the simulation of each avalanche starts with activating a single randomly chosen unit and ends when the network activity dies out.

Quasicritical networks exhibit apparent scale-free avalanche-size distributions with the expected MF-DP power-law exponent ($\tau \approx 1.5$, Fig. 3, top). In particular, small quasicritical networks with a finite interaction radius (e.g., $k = 3$ and $L = 8, 32$) can seemingly display finite-size scaling. The apparent scaling is more visible in networks with a larger connectivity radius k (Fig. 10). However, for sufficiently large system sizes, a characteristic scale becomes evident (cutoff stays independent of the system size when $L > L_{\mathrm{scale}}$). The characteristic scale in quasicritical networks becomes more apparent when compared to the avalanche-size distributions of critical networks (with the branching parameter found in Fig. 2) with the same size and topology (Fig. 3, bottom). In critical networks, power laws extend up to much larger sizes. For nearest-neighbor connectivity, the exponent shifts from mean-field values of $\tau \approx 1.5$ towards $\tau \approx 1.27$ as expected for the two-dimensional directed-percolation (2D-DP) universality class [27] (they approach $\tau = 1.27$ as $N \to \infty$, Fig. 11, Appendix D). Changing the network topology towards random or all-to-all connectivity brings the critical point closer to $m_{\mathrm{MF}} = 1$ (Fig. 2). Hence, in quasicritical networks with larger k, the characteristic scale of avalanche sizes increases and the exponent shifts towards 1.5 (Figs. 3 and 9).

In finite-size critical networks, increasing the connectivity radius or rewiring probability changes the critical exponents continuously from $\tau \approx 1.27$ (2D-DP) to $\tau \approx 1.5$ (MF-DP). While the theory of critical phenomena predicts that critical exponents differ between random and structured networks [19,27], here we find that these two mechanisms alter the exponents in different ways. For any finite connectivity radius k and no rewiring, in the limit of large network sizes, connections are short-ranged, and the dynamics belong to the 2D-DP universality class with the critical exponent of $\tau = 1.27$. At the same time, finite networks with large enough k are almost fully connected, showing exponents similar to MF-DP. Thus, for fixed k, the network size affects how close the system is to a fully connected system. The combination of these two factors leads to the true scaling exponent (known for the 2D-DP and MF-DP [27]) being visible only for very large avalanches ($S \gg 1$), which require very large system sizes ($L \to \infty$). In large networks, the power-law exponent will change slowly and continuously from MF-DP for relatively small avalanches

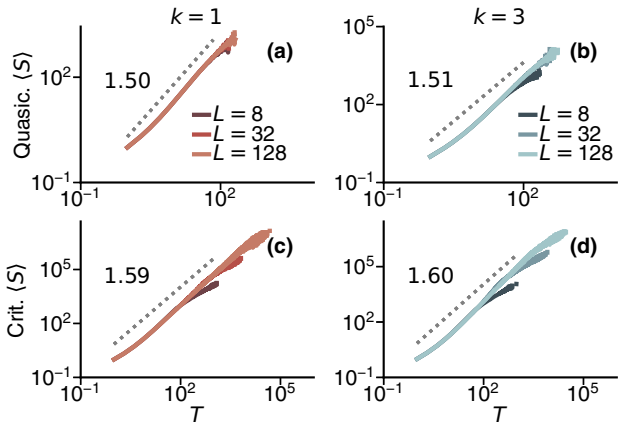

FIG. 4. The cracking noise relation for quasicritical (a), (b) and critical (c), (d) networks. For all networks, the exponent of the relation between avalanche durations and their mean avalanche size follows closely the value computed from $\frac{\alpha-1}{\tau-1}$ (gray dashed line). Distributions correspond to the same simulations as in Fig. 3. See Appendix F for more details.

to 2D-DP for large events (Fig. 12). For the rewiring case, in the thermodynamics limit, for any $p_{\mathrm{rw}} > 0$, the system was shown to exhibit MF-DP exponents [34,35]. However, for a finite network size, the avalanche-size exponent is smaller than $\tau = 1.5$ and approaches the MF-DP value with increasing p_{rw} (Fig. 13). For a close-to-small-world network structure ($p_{\mathrm{rw}} = 0.01$), the MF-DP exponents are only visible for large avalanches obtained from very large system sizes (10^6). As networks become closer to a random structure ($p_{\mathrm{rw}} \geqslant 0.1$), the MF-DP exponents can also be retrieved in smaller networks. Therefore, rewiring and increasing the connectivity radius change the exponents in distinct ways, particularly in the case of finite networks that are relevant for interpreting empirical observations, for example from neural recordings. The difference between the two cases can be understood in how the effective dimension of the network changes with increasing k or p_{rw} (Appendix E).

Our results suggest that observing the power-law-like behavior in avalanche-size distributions from *small* networks cannot be a reliable signature of criticality even when distributions scale with the system size. Close to criticality, subcritical systems can exhibit apparent scale-free behavior, where the characteristic scale is only uncovered in the limit of large systems. In addition to the finite-size scaling of avalanche-size distributions, the cracking noise relation, i.e., the relation between exponents of avalanche-size and -duration distributions, is commonly used as a measure for disentangling critical and noncritical dynamics [36–38]. For this purpose, we computed the avalanche duration distributions for critical and quasicritical networks (Fig. 14) and estimated their corresponding avalanche-size and avalanche-duration exponents (Fig. 4). We find that the crackling noise relation is fulfilled even for the quasicritical networks, despite deviations of exponents from the expected values in MF-DP (more details are in Appendix F). Hence, it is challenging to disentangle critical and quasicritical dynamics in small networks solely based on empirical measurements.

IV. COALESCENCE IN MEAN-FIELD AND STRUCTURED NETWORKS

We next investigate the mechanisms underlying the differences between the avalanche dynamics in mean-field and structured networks. Due to locally structured connectivity in our model, each unit i can be activated by multiple sources at the same time, generating *coalescence* [Fig. 6(a)]. We define the local coalescence as the number of sources that activated unit i minus 1. Let $n_{i,t}$ be a number of active neighbors of unit i at time t. The local coalescence of unit i at time t is a random variable

$$C_{i,t} = \max\left(0, \sum_{k=1}^{n_{i,t}} b_k + s_i b_0 - 1\right), \tag{5}$$

where $b_k \sim \mathrm{Bernoulli}(p_{\mathrm{r}})$ and $b_0 \sim \mathrm{Bernoulli}(p_{\mathrm{s}})$. Let A_t be the number of active units in the network at time t. Then, the average normalized network coalescence $C(A)$ is given by

$$C(A) = \frac{1}{A}\left\langle \sum_{i=1}^{L\times L} C_{i,t} \middle| A_t = A \right\rangle, \tag{6}$$

where the average is taken over all t with $A_t = A$. The network coalescence can be computed analytically [17] for a branching network with random or all-to-all connectivity as

$$m_{\mathrm{eff}}(A) = \frac{N}{A}\left(1 - \left(1 - \frac{m}{A}\right)^A\right). \tag{7}$$

However, the analytical determination of coalescence for finite-dimensional topologies (e.g., structured networks) requires renormalization-group approaches relying on the precise knowledge of the system dimension [19,20]. This approach becomes especially difficult when dealing with finite system sizes. Hence, we use a simulation-based approach to estimate the coalescence from the activity of structured networks. We can estimate the effective branching parameter for the given number of active units A from the simulated activity as

$$m_{\mathrm{eff}}(A) = m - C(A). \tag{8}$$

For these analyses, the simulations are performed in the timescale-separated regime.

Due to coalescence, the effective branching parameter is smaller than the local branching parameter and depends on the number of active units and network topology. For small A, $m_{\mathrm{eff}}(A)$ is closer to m, but with an increasing number of active units, the coalescence also increases, leading to larger deviations of $m_{\mathrm{eff}}(A)$ from m [Figs. 6(b) and 6(c), left]. Therefore, a larger local branching parameter is required to have an effective branching parameter close to 1 that creates critical dynamics. When increasing the connectivity radius or the rewiring probability, the activity can spread to a broader range of units. With increasing rewiring probability, neighboring units are less likely to be connected, reducing the coalescence. With increasing the connectivity radius while keeping the local branching parameter constant, the strength of individual connections reduces, and coalescence for every unit scales as a power of the connection strength. Thus, by increasing the connectivity radius or the rewiring probability, the coalescence

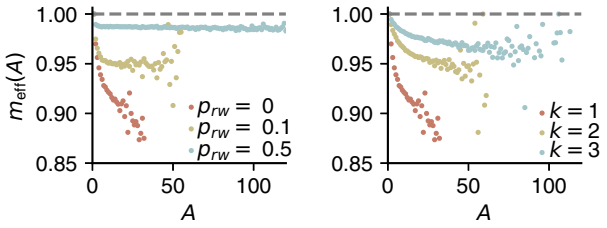

FIG. 5. Dependence of effective branching parameter m_{eff} on network topology. With increasing rewiring probability p_{rw} (left, $k = 1$) or connectivity radius k (right, $p_{\text{rw}} = 0$), the coalescence decreases and the effective branching parameter m_{eff} becomes closer to the local branching parameter m (dashed line). $L = 128$.

decreases, and the $m_{\text{eff}}(A)$ becomes closer to m (Fig. 5, consistent with Fig. 2). However, similar to the critical exponents, the two mechanisms affect the network coalescence in distinct ways (Fig. 5): the coalescence grows (i.e., m_{eff} reduces) with the number of active units we could observe in the case of increasing connectivity radius, whereas for increasing rewiring probability, it saturates at a certain level depending on p_{rw} and stops changing with the number of active units above a certain level $A > A_{\text{min}}$. For both mechanisms, the convergence to m as $k \to L/2$ or $p_{\text{rw}} \to 1$ is only realized at the thermodynamic limit; thus, for a finite-size mean-field network, we would still observe some level of coalescence depending on the network size and the number of active units [Eq. (7)].

To capture the impact of coalescence on avalanche-size distributions, we simulated an equivalent adaptive branching process for each network model. To generate this process, we first find $m_{\text{eff}}(A)$ from a long simulation of the network model using Eq. (8). Then we define the adaptive branching process as a Markov process $\tilde{A}(t)$, where each of the ancestors can generate a binomially distributed number of offsprings z_i,

$$\tilde{A}_{t+1} = \sum_{i=1}^{\tilde{A}_t} z_i(\tilde{A}_t), \quad z_i \sim \text{Bin}\left(n, \frac{m_{\text{eff}}(\tilde{A}_t)}{n}\right). \tag{9}$$

n is the maximal number of offsprings for one ancestor. For simulations, we took $n = 8$.

Avalanche-size distributions of the adaptive branching processes well approximate the shape of corresponding distributions from the network dynamics [Figs. 6(b) and 6(c), right]. In the quasicritical network, distributions for the branching process and the network are completely overlapping. The adaptive branching process for the critical network captures most of the distribution and the power-law exponent. The mismatch in the tail can be generated by a too scarce sampling of large avalanches to estimate the correct effective branching parameter. Overall, despite the large variability in the network coalescence, the average effective branching parameter is sufficient to predict the shape and exponent of avalanche-size distributions in the structured networks.

V. DISCUSSION

Our results indicate that differences between the dynamics in structured and mean-field networks arise from the coalescence created by the local network interactions. Increasing

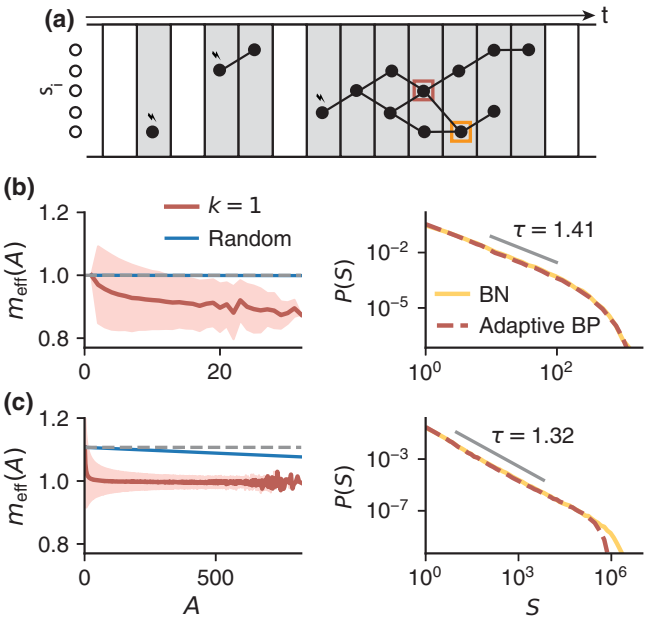

FIG. 6. The adaptive branching process captures the shape of avalanche-size distributions in structured networks ($k = 1, L = 128$). (a) Avalanches (gray frames: each row, one unit; filled circles, active units) are separated by quiescence (white frames). Two types of coalescence: simultaneous activation by multiple neighbors (brown square) or by self-excitation and a neighbor (orange square). Each avalanche starts with the external input (black lightning bolts). The arrow indicates time. (b), (left) Effective branching parameter $m_{\text{eff}}(A)$ (brown and blue lines) in the quasicritical ($m = 1$) network as a function of the number of active units [A, Eq. (8)] deviates from the local branching parameter (dashed line). The deviations are larger for the local network ($k = 1$, brown) than the random network (blue, computed analytically). Small deviations for the random network are not visible in the figure. Shading: ± 1 s.d. of $C_{i,t}|A$ for the local network. (b), (right) The adaptive branching process (dashed brown) has the same avalanche-size distribution $P(S)$ as the structured branching network (BN, yellow line). Gray line: power-law fit with exponent τ. (c) Same as (b) for the critical network ($m = 1.109$). Deviations are more visible for the random network than in (b) since the number of active units A reaches higher values.

the connectivity radius or rewiring probability reduces the coalescence, and the network dynamics becomes more similar to the conventional branching process. The dependence of coalescence on network topology is in line with the reduction of the effective branching parameter with increasing network clustering [39]. Although it has been shown that certain connectivity structures [35,40,41] can facilitate critical dynamics, there was no systematic study of the impact of varying coalescence on avalanche distributions.

One of the common methods for estimating the location of the critical point in branching networks is using the largest eigenvalue of the connectivity matrix instead of the local branching parameter. This method is particularly useful for networks with inhibition [42] or depressing synapses [43]. However, changing the network structure by either increasing the connectivity radius or rewiring only affects the intermediate eigenvalues, and the largest eigenvalue is always equal to the local branching parameter. Therefore, the computation

of coalescence from network activity is required to determine the location of the critical point.

We demonstrated how a range of scaling exponents can arise from the network structure and be misinterpreted in finite-size networks. Specifically, quasicritical avalanche-size distributions can appear as critical in small networks. We further showed that similar behavior could be observed in the activity of subsampled units from a larger network (Appendix G). This is particularly significant for interpreting observations from finite-size neural recordings, where the number of recorded neurons or electrodes is often treated as a proxy for system size, and finite-size scaling is assessed by downsampling the recorded neurons (electrodes) [44]. The precise number of neurons involved in the dynamics and their interaction radius is often unknown, and approaches that can recover the dynamical regime from subsampled networks [45] so far mainly rely on the downsampling of available data.

Despite the common assumption that neural dynamics operate close to a critical point belonging to the MF-DP universality class, estimated critical exponents from neural activity vary. Observed exponents take values smaller [3,46] or larger [37,38] than MF-DP exponents, depending on the methodology applied for avalanche detection and the type of neural data (e.g., spikes, local-field potentials, calcium imaging, *in vivo* or *in vitro*, different behavioral states). We showed that increasing the connectivity radius and rewiring can give rise to different critical exponents between 2D-DP and MF-DP values by affecting the dimensionality of finite-size (or subsampled) networks in distinct ways. These results indicate how within the same universality class, brain networks' architecture might affect measured critical exponents from subsampled neural data. Diverse critical exponents can also arise due to external input [47] and the absence of separation of timescales in measured avalanches [48,49]. At the same time, distinct critical exponents might reflect other universality classes, such as a phase transition at the onset of collective oscillations [37,50,51]. Therefore, a combination of different mechanisms might be involved in shaping statistics of observed neural avalanches. A more thorough evaluation of experimental findings (considering various methods and neural data modalities) and theoretical investigations on interactions between different mechanisms can help better understand the underlying mechanisms for variable exponents in neural activity.

Codes for simulating the network models and reproducing the analyses are publicly available on GitHub [52].

ACKNOWLEDGMENTS

This work was supported by a Sofja Kovalevskaja Award from the Alexander von Humboldt Foundation, endowed by the Federal Ministry of Education and Research (R.Z., V.B., A.L.), the SMARTSTART2 program provided by Bernstein Center for Computational Neuroscience and Volkswagen Foundation (R.Z.), NIH grants RF1DA055666 (T.A.E.), and the Sloan Research Fellowship from the Alfred Sloan Foundation (T.A.E.). Simulations were performed with assistance from the NIH Grant No. S10OD028632-01. V.B. acknowledges the Spanish Ministry and Agencia Estatal de investigación (AEI) through Project I+D+i Ref. PID2020-113681GB-I00, financed by MICIN/AEI/10.13039/501100011033. We acknowledge the support from the BMBF through the Tübingen AI Center (FKZ: 01IS18039A) and International Max Planck Research School for the Mechanisms of Mental Function and Dysfunction (IMPRS-MMFD). We thank Miguel A. Muñoz and Giorgio Nicoletti for valuable discussions.

APPENDIX A: ESTIMATING THE CRITICAL BRANCHING PARAMETER USING PHASE DIAGRAMS

The critical branching parameter has been estimated by computing the location of the maximum activity susceptibility. Estimation of the critical transition's location for absorbing-active phase transitions is complicated since at criticality the stationary state is the absorbing one. Hence, at criticality one is forced to make statistics over pseudostationary states [19,20]. Phase diagrams are computed by randomly activating 15% of the system size and letting the simulations run for a fixed time $t_{\text{sim}} = 10^5$, which is set to be as large as possible. The initial percentage is set to a low value so that the simulation thermalizes fast, but at the same time large enough to avoid supercritical simulations falling into the absorbing state due to statistical fluctuations. This is crucial since close to criticality relaxation times become very long. The first two moments of the total particle density are measured during this time, allowing enough time separation between measurements to avoid any correlation bias. These allow us to obtain the average density $\langle\rho\rangle$ and the susceptibility $\chi(\rho) = \sqrt{N}(\langle\rho^2\rangle - \langle\rho\rangle^2)$. If the simulation falls into the absorbing state before reaching t_{sim}, measurements are discarded (setting mean and variance equal to zero) and the procedure starts again, so the results are averaged only over runs that survived. Near criticality, even when activity eventually falls to zero, a small density is still able to produce avalanches at any time, so there is always a nonvanishing probability of observing any amount of activity at t_{sim}, which grows with system size. This is different from the active phase, where activity can be arbitrarily small, but fluctuates around its mean value.

Then, one computes the average density and its susceptibility, looking for the largest susceptibility to have an estimation for the location of the critical transition. To get a better estimation for the critical control parameter, the usual technique is finite-size scaling, computing the average density for survived runs for long simulation durations and increasing sizes. Criticality fulfills power-law decay of the density ρ^{-N}, while the active phase saturates and the subcritical decays exponentially. However, this method requires very long simulation times, large system sizes, and many runs in order to have good statistics over the survival ones, which in our model lead to very long computation times [54].

Finally, in contrast with continuous-time models, here we use a nonlinear probabilistic model simulated in discrete times, leading to a more complex behavior for the transition rates than in the classical contact process. Hence, the exact location of the critical point needs a very accurate determination of the recurrent probability p_r, as demonstrated in Appendix C.

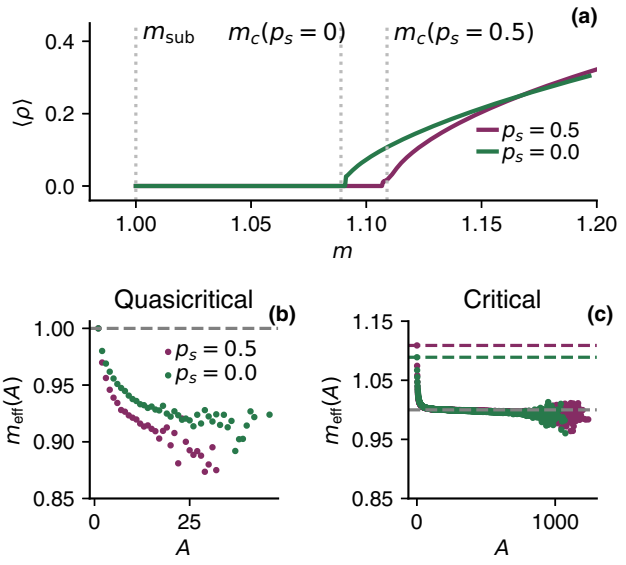

FIG. 7. Dependence of the effective branching parameter on the self-excitation probability. (a) Location of the critical point moves slightly with p_s. (b) In quasicritical networks with the same local branching parameter ($m = 1$, dashed line), different values of p_s give rise to a different amount of coalescence in network activity and m_{eff}. (c) At the critical point ($m_c = 1.109$ for $p_s = 0.5$ to $m_c = 1.089$ for $p_s = 0$, colored dashed lines), the amount of coalescence is independent of p_s.

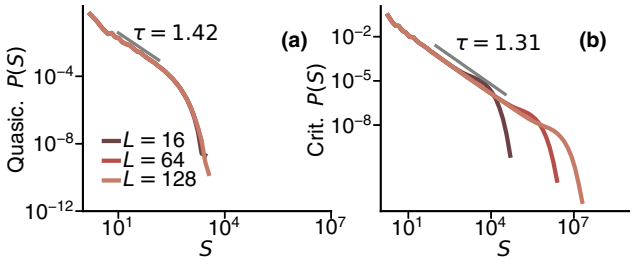

FIG. 8. Avalanche-size distribution in the absence of self-excitation. For the critical network, the cutoff location of the avalanche-size distribution shifts with system size (L^2) (right), while quasicritical avalanches exhibit a characteristic scale for large systems (left). Gray lines indicate the fitted power-law distribution with the exponent τ. $L = 128$, $k = 1$ (similar results to Fig. 3).

APPENDIX B: DEPENDENCE OF CRITICAL PARAMETERS ON THE SELF-EXCITATION PROBABILITY

The dynamics of the networks depends on two parameters, self-excitation probability p_s and recurrent-excitation probability p_r, which relate to each other through local branching parameter $m = p_s + np_r$, where n is the number of neighbors of each unit. In the main text, we set a fixed value of $p_s = 0.5$ and described the system's dynamics as a function of local branching parameter m that is similar to fixing the recovery rate in the contact process [19].

We show that while the location of the critical point can move slightly with p_s, at the critical point, the macroscopic properties of dynamics are independent of p_s. In the absence of self-excitation ($p_s = 0$), the location of the critical point determined by the phase diagrams moves slightly toward the mean-field value [i.e., from $m_c = 1.109$ for $p_s = 0.5$ to $m_c = 1.089$ for $p_s = 0$, for the structured network with $k = 1$ and $L = 128$, Fig. 7(a)], suggesting that $p_s = 0.5$ generates a larger amount of coalescence. We can demonstrate this effect by comparing the amount of coalescence for quasicritical networks. With fixed $m = 1$, the network with $p_s = 0.5$ exhibits larger coalescence (i.e., smaller m_{eff}) than the network with $p_s = 0$ [Fig. 7(b)]. However, at the critical point for each network ($m_c = 1.109$ for $p_s = 0.5$ to $m_c = 1.089$ for $p_s = 0$), the amount of coalescence is independent of p_s [Fig. 7(c)], since it is compensated by different values of m. This difference between critical and quasicritical networks is due to the fact that quasicritical $m = 1$ is closer to actual criticality for $p_s = 0$. Moreover, the avalanche-size distributions and their exponent for $p_s = 0$ [Fig. 7(c)] follow similar patterns and values to those for $p_s = 0.5$ (Fig. 3) in both critical and quasicritical networks. These results suggest that the amount of p_s does not create significant differences in macroscopic dynamics, although it affects microscopic properties of dynamics [10,31]. Additionally, we report the avalanche-size distributions for the case $p_s = 0$ in Fig. 8, which are similar to the ones reported for $p_s = 0.5$.

APPENDIX C: CONTINUOUS MEAN-FIELD APPROACH

It is possible to demonstrate that our discrete model has a second-order phase transition in the mean-field, and to obtain an exact relationship between the probabilities p_s, p_r, and the branching ratio m. To do so, we proceed in the following way: first, the discrete probabilities are written as a continuous Markov process, from which it is possible to derive a Master equation to apply our formalism; second, one performs a Kramers-Moyal expansion of the master equation, from which it is possible to identify a Langevin dynamics for the density of active particles; finally, this equation is expanded near the absorbing state and mapped to the "normal form" of the contact process.

First, under the mean-field approach, the transition probability of a single node becoming active is given by

$$p(0 \to 1) = 1 - (1 - p_r)^x \equiv F(x), \tag{C1}$$

where $x = A/N$ is the particle density of the system, which is an intensive variable, well-defined in the thermodynamic limit. Thus, the transition probability matrix between states is given by

$$\hat{P}(\Delta t) = \begin{pmatrix} 1 - F(x) & F(x) \\ (1 - p_s)[1 - F(x)] & p_s + (1 - p_s)F(x) \end{pmatrix}. \tag{C2}$$

In a continuous time model described by the Markov transition matrix $\hat{Q}$, the probability that a transition took place during the time step Δt is given by $\hat{P} = \exp(\hat{Q}\Delta t)$. Equating both allows us to find the Markov transition rates [31],

$$\omega(0 \to 1) = -\frac{F(x)}{\Delta t\{1 - p_s[1 - F(x)]\}} \log[(1 - p_s)F(x)], \tag{C3}$$

$$\omega(1 \to 0) = -\frac{[1 - F(x)](1 - p_s)}{\Delta t\{1 - p_s[1 - F(x)]\}} \log[(1 - p_s)F(x)]. \tag{C4}$$

In the mean-field, since all the particles are identical, the probability of increasing the activity by one particle is given by the probability of picking an empty site and performing a transition up. Conversely, the probability of decreasing activity is given by the probability of picking an active particle and transitioning down, i.e.,

$$\Omega(x \to x + \Delta x) = (1 - x)\omega(0 \to 1), \tag{C5}$$

$$\Omega(x \to x - \Delta x) = x\omega(1 \to 0). \tag{C6}$$

An interesting theoretical note is to realize that the nonlinear activation rate will yield arbitrary powers of the density x when it is Taylor-expanded. This can be interpreted as having n-body interactions, since in simple models with linear rates transitions involving n bodies have rates proportional to x^n (the contact process, for example, only involves up to a quadratic term). This is a direct consequence of coalescence, and in practice it means that in the discrete model a particle in contact with two active neighbors can be activated by either one of those or by the effect of both acting together. So, if rates are Taylor-expanded around the absorbing state $x = 0$,

$$\Omega(x \to x \pm \Delta x) = \sum_{k=1}^{+\infty} \lambda_k^{\pm} x^k, \tag{C7}$$

then the rate at which a particle in contact with two active neighbors activates is given by $2\lambda_1^+ + \lambda_2^+$.

Once the global rates have been identified, one can write a Master equation and expand it using the Kramers-Moyal approximation. Since this is a standard procedure, we will skip the technical details, redirecting the reader instead to a classic textbook on the subject [55]. One can then show that a Langevin equation for the density of active particles is given by

$$\dot{x} = \Omega(x \to x + \Delta x) - \Omega(x \to x - \Delta x) + \frac{1}{\sqrt{N}}\sqrt{\Omega(x \to x + \Delta x) + \Omega(x \to x - \Delta x)}\xi(t), \tag{C8}$$

where $\xi(t)$ is a Gaussian, δ-correlated white noise. Finally, the rates are Taylor-expanded. Following the Landau-Ginzburg theory of critical phenomena, the critical properties of the transition are to be controlled by the first term that becomes always negative (hence, controlling saturation) [33]. It is possible to show that it is sufficient to expand the equation up to second order,

$$\dot{x} = a_1 x - a_2 x^2 + \sigma\sqrt{x}\xi(t), \tag{C9}$$

where

$$a_1 = \frac{1 - p_s - \log(1 - p_r)}{1 - p_s} \log p_s, \tag{C10}$$

$$a_2 = -\frac{\log(1 - p_r)}{2(1 - p_s)^2}[2(1 - p_s)^2 + \log(1 - p_r)[2(1 - p_s) + (1 + p_s)\log p_s]]. \tag{C11}$$

One can demonstrate that $a_2 > 0$ always by direct plotting, or more elegantly, by demonstrating that the function is monotonously increasing and its minimum is positive [56]. The critical point happens when the linear term (the "mass") vanishes, which happens at $-\log(1 - p_r) = 1 - p_s^*$. Finally, it is possible to evaluate the branching ratio of the continuous model, knowing that in the contact process we have $a_1 = (1 - m)a_2$. The actual relation between the branching ratio and the probabilities is then

$$m = \frac{\log(1 - p_r)[2(1 - p_r)^2 + \log(1 - p_r)[2(1 - p_s) + (1 + p_s)\log p_s]]}{2(1 - p_s)\log(1 - p_r)(1 - p_s + \log p_s) + 2(1 - p_s)^2 \log p_s + \log^2(1 - p_r)[2(1 - p_s) + (1 + p_s)\log p_s]}, \tag{C12}$$

which reduces to $m = 1$ when the critical p_r^* is set. One can then see that small changes to the probabilities translate into nonlinear changes to the branching ratio, which makes the model extremely sensitive to the choice of probability when trying to locate the critical point, and makes it difficult to find a clean scaling for the avalanches for the structured networks. In the mean-field case, however, the critical point can be found exactly, and avalanches with the expected exponents are found, as shown in Fig. 9.

APPENDIX D: ESTIMATING THE POWER-LAW EXPONENT FROM THE AVALANCHE-SIZE DISTRIBUTIONS

We compute the avalanche-size distributions in the separated timescale regime ($p_{\text{ext}} = 0$). Each avalanche starts with a single active unit and ends when the whole network activity dies out. We define the size of an avalanche as the total number of units activated during the avalanche. To obtain the avalanche-size distributions, we simulated 10^7 avalanches for each network. For coalescence analysis and the adaptive branching process, we simulated 10^5 avalanches.

At the critical point, the size of avalanches S follows a power-law distribution. We estimate the power-law exponent τ by fitting the avalanche-size distribution with a discrete and truncated power-law distribution as [57,58]

$$P(S) = \frac{S^{-\tau}}{\zeta(\tau, S_{\min}) - \zeta(\tau, S_{\max})}. \tag{D1}$$

Here, $S_{\min}$ and $S_{\max}$ are, respectively, the minimum and maximum avalanche size considered for fitting, and $\zeta(\tau, S)$ is the Hurwitz zeta function defined as

$$\zeta(\tau, S) = \sum_{n=0}^{\infty}(n + S)^{-\tau}. \tag{D2}$$

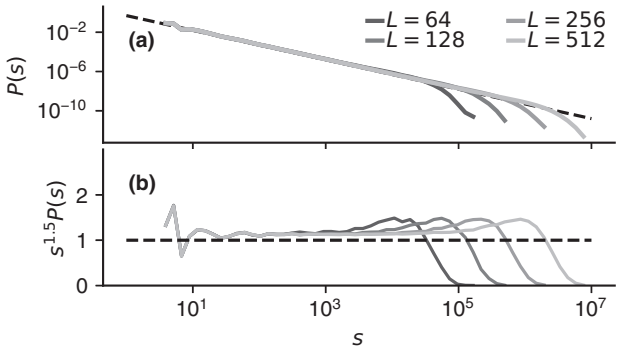

FIG. 9. Scaling of mean-field avalanches. (a) Finite-size scaling of the mean-field avalanche size distribution for different system sizes at criticality, $p_s = 0.5$ and $p_r = 0.393\,469$. Theoretical scaling $P(s) \sim s^{-3/2}$ is displayed with a discontinuous line. (b) The same distribution, multiplied by $s^{3/2}$ to display it as a horizontal line, in order to ease the visual inspection of the correct scaling. The distributions were obtained with 10^7 avalanches.

We find the optimal value of τ using the maximum-likelihood estimation (MLE) with a grid search. For the fits, we set $S_{\min} = 10$ and $S_{\max}$ to the 96th percentile of the distribution.

We found that the estimated exponents depend on the network topology (Fig. 3). In particular, Fig. 10 shows the avalanche-size distribution for the $k = 5$ quasicritical network, which has an exponent closer to 1.5 than the low-k cases shown in the main text [Figs. 3(b) and 3(c)]. This is because the critical local branching parameter m_c approaches 1 as k increases for fixed N, meaning that the quasicritical network becomes closer to actual criticality with a larger k. We also observe an apparent shift of the cutoff with the system size (that will disappear for even larger networks). Therefore, finite structured quasicritical networks can appear critical when the connectivity radius is large enough (Fig. 10) or the network size is small [Figs. 3(b) and 3(c)].

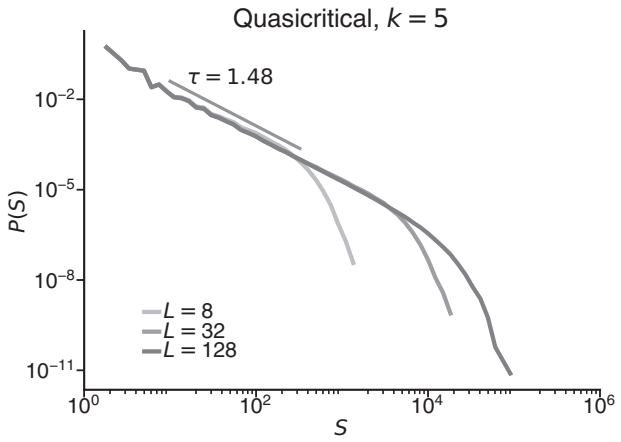

FIG. 10. Quasicritical ($m = 1$) avalanche size distribution for $k = 5$. At finite sizes, avalanches for larger k present seemingly critical distributions, with a cutoff shifting with the system size, and the mean-field exponent close to 1.5.

APPENDIX E: INCREASING RADIUS VERSUS REWIRING: THE EFFECT OF DIMENSIONALITY

We argued in the main text that increasing both the connectivity radius and rewiring probability brings the critical exponents closer to the mean-field ones, while exhibiting different behaviors in the thermodynamic limit behavior. In this Appendix, we show how the topology modulates the dimension of the system. Dimensionality plays an essential role in the theory of critical phenomena [19,20,33]. One can see that in the increasing radius case, the critical exponents must be continuously varying based on the following argument: for any arbitrarily large size N, one can always take a large enough k to make the network almost fully connected, displaying avalanches with near-mean-field scaling, up to size $S_{\max}(N)$, which depends on N. For $S > S_{\max}$, $P(S)$ is no longer described by a power-law due to finite-size effects. Now, if N is further increased to $N' \geqslant N$, the distribution up to $S_{\max}(N)$ must be exactly the same as before, since by hypothesis we assumed all boundary effects are taking place for larger sizes. At the same time, we know that for a critical system we have finite-size scaling, so the scale-free distribution now must hold for $S_{\max}(N) < S < S_{\max}(N')$. But if $N' \to \infty$, then $k \ll N'$ again, making the network structured so the exponent of this new scaling should be different from the one presented until $S_{\max}$.

One could argue that by increasing N', the network's effective dimensionality is reduced to $d = 2$. In fact, the classical definition of dimension is that the mass encompassed in the ball of radius r scales as $M \sim r^d$. When this scaling relationship is not present, it is assumed that the dimension is not well defined (i.e., $d \to \infty$). One can naively generalize the dimension definition to networks by letting r be the distance between nodes, and $\langle M \rangle$ the average total number of nodes at a distance less than or equal to r. For the structured lattice, it is clear that $M \sim r^2$, as long as the network is infinite. If the network is finite, then there is a distance $r^*(k, N)$ such that $\langle M(r \geqslant r^*) \rangle = N$. Hence, the network appears to be 2D

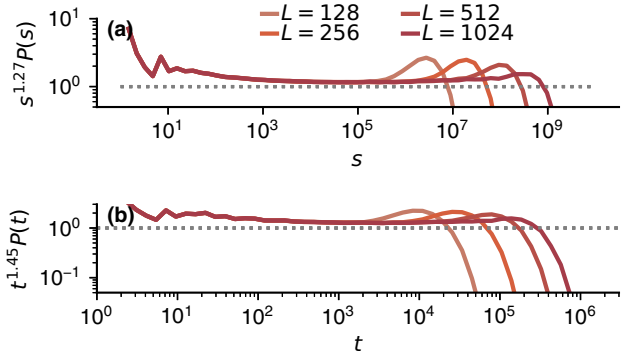

FIG. 11. True scaling of the system in the thermodynamic limit at criticality for $k = 1$. Avalanches of very large system sizes are plotted to show that as $N \to +\infty$ the expected 2D directed percolation exponents ($\tau = 1.27$ for sizes and $\alpha = 1.45$ for duration [27]) are recovered. Avalanche-size (a) and duration distributions (b), scaled to render distributions horizontally based on the 2D directed percolation exponents (dashed lines) to appreciate the possible deviations from the theoretical scaling in detail. Parameters, $p_s = 0$, $p_r = 1.089\,75$, and 10^7 avalanches.

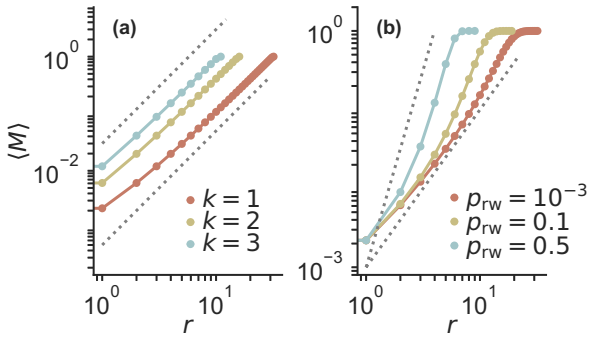

FIG. 12. Dimensionality estimation based on topology. (a) Scaling of $\langle M \rangle$—the average total number of nodes at a distance less than or equal to r—as r increases for different connectivity radii k. Notice that the slope is always the same, $d = 2$, marked with dashed black lines, but as k grows, the saturation appears in smaller k. (b) Same as (a) for increasing rewiring probability p_{rw}. The saturation point changes due to different slopes, increasing with p_{rw}. Dimensions $d = 2$ and 5 are indicated for reference (dashed back lines).

for $r \leqslant r^*$. In practice, even for low values of k one needs huge sizes to see the 2D network scaling. Figure 11 shows that indeed the structured system with $k = 1$ relaxes to the critical exponent of the directed percolation universality class, $\tau \approx 1.27$. However, notice that the scaling is not clear until reaching very large system sizes.

For the case of rewiring, long-ranged connections allow connecting any arbitrary pair of nodes in a small number of steps, making again $\langle M \rangle = N$ for small distances. However, in this case, as rewiring probability p_{rw} is increased, so does the slope of $\langle M(r) \rangle$. In this case, the scaling relation is only lost if the network is completely random, since this is the only case in which any two nodes could be possibly connected at a finite distance. Any small amount of structure will make certain nodes infinitely separated, allowing them to fulfill the scaling for $\langle M \rangle$. Figure 12 illustrates the differences in the function $\langle M \rangle$ between the increasing radius and the rewiring cases. Similar to the case of increasing radius, the true scaling exponents can only be observed at the thermodynamic limit (Fig. 13).

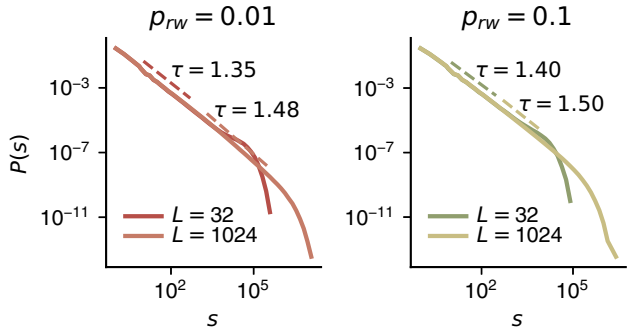

FIG. 13. Dependence of avalanche-size distributions on rewiring probability p_{rw} and system size L^2. For a large system size ($L = 1024$), the exponent of the avalanche-size distribution (shown as dashed lines) converges to 1.5, the exponent of MF-DP. However, for smaller systems ($L = 32$), the exponent largely depends on the rewiring probability.

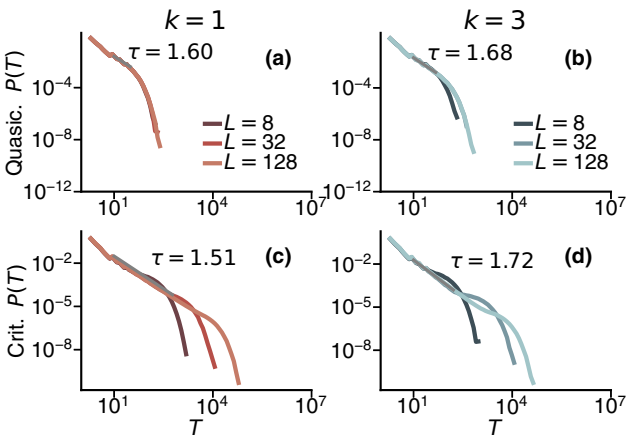

FIG. 14. Avalanche-duration distributions for quasicritical (a), (b) and critical (c), (d) networks, for $k = 1$ (a), (c) and $k = 3$ (b), (d). Distributions correspond to the same simulations as in Fig. 3.

Differences between increasing the connectivity radius and the rewiring probability can also be observed when measuring the effective branching parameter from the dynamics of these two network types (Fig. 5). Due to the coalescence, the effective branching parameter is smaller than the local branching parameter.

APPENDIX F: VALIDITY OF THE CRACKLING NOISE RELATION

The size and duration of the avalanches are not independent. The relation between avalanche duration T and mean avalanche size $\langle S \rangle$ for this given duration is characterized by $\langle S \rangle \sim T^{\gamma}$. At the critical point, exponent γ is directly related to the exponents of avalanche-size and -duration distributions. This relationship, known as the crackling noise relation, can be obtained for the directed percolation universality class [27,37] as

$$\frac{\alpha - 1}{\tau - 1} = \gamma. \tag{F1}$$

The absence of a cracking noise relation between empirical exponents can be used as a signature for the absence of critical dynamics [59].

We show that the cracking noise relation holds in both critical and quasicritical networks (Fig. 4). Using the same simulations as in Fig. 3, we compute the avalanche-size and -duration distributions and compare the relationship between their exponents to the cracking noise relation. Similar to the avalanche sizes (Fig. 3), the exponents of avalanche-duration distribution change with network topology (Fig. 14). For the critical structured network ($k = 1$), $\alpha \sim 1.5$ is close the expected exponent of 1.45 for 2D-DP, while the exponent for quasicritical networks has larger deviations from the expected value of 2 for the MF-DP. Nevertheless, the cracking noise relation holds for both critical and quasicritical networks, suggesting that just fulfilling this relation might not be a good indicator of criticality: as long as both sizes and duration distributions are scale-free, it is possible to fulfill the relation, even at a quasicritical regime. The only hint that suggests that the system is not critical is the deviation of the duration ex-

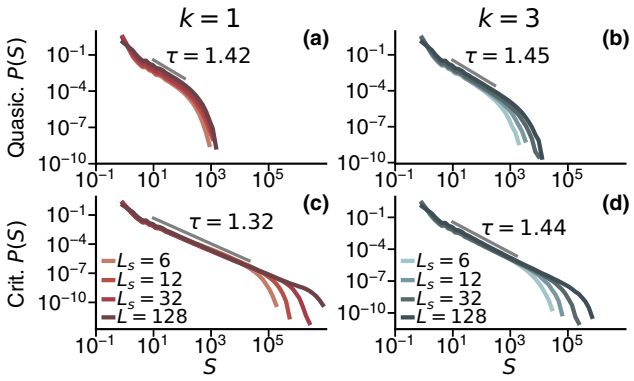

FIG. 15. Impact of windowed subsampling on quasicritical and critical dynamics. For each network, avalanche-size (S) distributions are measured from the subsampled window with size L_s^2. For critical systems, the cutoff location of the avalanche-size distribution shifts with window size. Quasicritical avalanches follow a power law up to a cutoff that is scaling with the window size for small systems. Gray lines indicate the fitted power-law distribution with exponent τ. Simulations are performed for $k = 1$ and 3 with the full system size of side $L = 128$. Subsample sizes are chosen so the subsampled window 6×6 (covers 5%), 12×12 (10%), and 32×32 (20% percent of the total area).

ponents (and hence the γ exponent) from the expected values in DP. However, in experimental data, it is difficult to assess if different exponents come from a quasicritical spreading process, or a critical system with a different universality class [37].

APPENDIX G: IMPACT OF SUBSAMPLING ON CRITICAL AND QUASICRITICAL DYNAMICS

In the main text, we showed how finite system size might hinder the dynamical state of a network, as quasicritical dynamics in small structured networks might appear as critical. Here we test whether the same behavior can be observed in subsampled units from a larger network, which is a more realistic setting for comparison for neuroscience experiments.

Motivated by multielectrode array recordings, we perform a windowed subsampling, where avalanches are recorded only in a window of size $L_s \times L_s$ out of the whole system [53]. We analyze the avalanche statistics within the subsampled window for critical and quasicritical networks. An avalanche starts with the activation of one of the subsampled units inside the window and ends when no active unit remains in the window. Notice that the activity might continue spreading in the network afterward. In the original network, this would count as a single avalanche, but in the subsampled network it might be observed as multiple avalanches.

We find that the effect of subsampling is similar to that of finite-size scaling. Quasicritical avalanche-size distributions follow an apparent power law with exponents close to MF-DP and their cutoff shifts with the number of subsampled units [Figs. 15(a) and 15(b)]. The subcritical nature of quasicritical dynamics is observed in comparison with critical networks that scale beyond the subsampled size and exhibit exponents close to 2D-DP [Figs. 15(c) and 15(d)]. Therefore, it is difficult to distinguish between critical and quasicritical dynamics from small subsampled recordings of neural activity.

[1] J. M. Beggs and D. Plenz, Neuronal avalanches in neocortical circuits, J. Neurosci. **23**, 11167 (2003).

[2] J. M. Beggs, The criticality hypothesis: how local cortical networks might optimize information processing, Philos. Trans. R. Soc. A **366**, 329 (2008).

[3] T. Petermann, T. C. Thiagarajan, M. A. Lebedev, M. A. L. Nicolelis, D. R. Chialvo, and D. Plenz, Spontaneous cortical activity in awake monkeys composed of neuronal avalanches, Proc. Natl. Acad. Sci. USA **106**, 15921 (2009).

[4] V. Pasquale, P. Massobrio, L. L. Bologna, M. Chiappalone, and S. Martinoia, Self-organization and neuronal avalanches in networks of dissociated cortical neurons, Neuroscience **153**, 1354 (2008).

[5] T. Bellay, A. Klaus, S. Seshadri, and D. Plenz, Irregular spiking of pyramidal neurons organizes as scale-invariant neuronal avalanches in the awake state, eLife **4**, e07224 (2015).

[6] M. A. Muñoz, *Colloquium:* Criticality and dynamical scaling in living systems, Rev. Mod. Phys. **90**, 031001 (2018).

[7] S. Safavi, M. Chalk, N. Logothetis, and A. Levina, Signatures of criticality in efficient coding networks, bioRxiv 2023 (2023).

[8] O. Kinouchi and M. Copelli, Optimal dynamical range of excitable networks at criticality, Nat. Phys. **2**, 348 (2006).

[9] J. Zierenberg, J. Wilting, V. Priesemann, and A. Levina, Tailored ensembles of neural networks optimize sensitivity to stimulus statistics, Phys. Rev. Res. **2**, 013115 (2020).

[10] R. Zeraati, Y.-L. Shi, N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore, A. Levina, and T. A. Engel, Intrinsic timescales in the visual cortex change with selective attention and reflect spatial connectivity, Nat. Commun. **14**, 1858 (2023).

[11] N. Bertschinger and T. Natschläger, Real-time computation at the edge of chaos in recurrent neural networks, Neural Comput. **16**, 1413 (2004).

[12] N. Tomen, D. Rotermund, and U. Ernst, Marginally subcritical dynamics explain enhanced stimulus discriminability under attention, Front. Syst. Neurosci. **8**, 151 (2014).

[13] R. Zeraati, V. Priesemann, and A. Levina, Self-organization toward criticality by synaptic plasticity, Front. Phys. **9**, 619661 (2021).

[14] V. Buendía, S. di Santo, J. A. Bonachela, and M. A. Muñoz, Feedback mechanisms for self-organization to the edge of a phase transition, Front. Phys. **8**, 333 (2020).

[15] S.-S. Poil, A. van Ooyen, and K. Linkenkaer-Hansen, Avalanche dynamics of human brain oscillations: Relation to critical branching processes and temporal correlations, Hum. Brain Mapp. **29**, 770 (2008).

[16] J. M. Palva, A. Zhigalov, J. Hirvonen, O. Korhonen, K. Linkenkaer-Hansen, and S. Palva, Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws, Proc. Natl. Acad. Sci. USA **110**, 3585 (2013).

[17] J. Zierenberg, J. Wilting, V. Priesemann, and A. Levina, Description of spreading dynamics by microscopic network

models and macroscopic branching processes can differ due to coalescence, Phys. Rev. E **101**, 022301 (2020).
[18] T. E. Harris, *The Theory of Branching Processes* (Springer, Berlin, Heidelberg, 2002).
[19] H. Hinrichsen, Non-equilibrium critical phenomena and phase transitions into absorbing states, Adv. Phys. **49**, 815 (2000).
[20] M. Henkel, H. Hinrichsen, and S. Lübeck, *Non-Equilibrium Phase Transitions: Absorbing Phase Transitions*, Theoretical and Mathematical Physics (Springer, Berlin, 2008).
[21] V. B. Mountcastle, The columnar organization of the neocortex, Brain **120**, 701 (1997).
[22] M. F. Casanova and E. L. Casanova, The modular organization of the cerebral cortex: Evolutionary significance and possible links to neurodevelopmental conditions, J. Comp. Neurol. **527**, 1720 (2019).
[23] M. A. Smith and A. Kohn, Spatial and temporal scales of neuronal correlation in primary visual cortex, J. Neurosci. **28**, 12591 (2008).
[24] S. Safavi, A. Dwarakanath, V. Kapoor, J. Werner, N. G. Hatsopoulos, N. K. Logothetis, and T. I. Panagiotaropoulos, Nonmonotonic spatial structure of interneuronal correlations in prefrontal microcircuits, Proc. Natl. Acad. Sci. USA **115**, E3539 (2018).
[25] S. Safavi, T. I. Panagiotaropoulos, V. Kapoor, J. F. Ramirez-Villegas, N. K. Logothetis, and M. Besserve, Uncovering the organization of neural circuits with generalized phase locking analysis, PLoS Comput. Biol. **19**, e1010983 (2023).
[26] R. Rosenbaum, M. A. Smith, A. Kohn, J. E. Rubin, and B. Doiron, The spatial structure of correlated neuronal variability, Nat. Neurosci. **20**, 107 (2017).
[27] M. A. Muñoz, R. Dickman, A. Vespignani, and S. Zapperi, Avalanche and spreading exponents in systems with absorbing states, Phys. Rev. E **59**, 6175 (1999).
[28] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, Nature (London) **393**, 440 (1998).
[29] T. A. Engel, N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore, and K. Boahen, Selective modulation of cortical state during spatial attention, Science **354**, 1140 (2016).
[30] J. van Kempen, M. A. Gieselmann, M. Boyd, N. A. Steinmetz, T. Moore, T. A. Engel, and A. Thiele, Top-down coordination of local cortical state during selective attention, Neuron **109**, 894 (2021).
[31] Y.-L. Shi, R. Zeraati, A. Levina, and T. A. Engel, Spatial and temporal correlations in neural networks with structured connectivity, Phys. Rev. Res. **5**, 013005 (2023).
[32] C. Haldeman and J. M. Beggs, Critical branching captures activity in living neural networks and maximizes the number of metastable states, Phys. Rev. Lett. **94**, 058101 (2005).
[33] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group* (Oxford University Press, New York, 2001).
[34] R. Pazzini, O. Kinouchi, and A. A. Costa, Neuronal avalanches in watts-strogatz networks of stochastic spiking neurons, Phys. Rev. E **104**, 014137 (2021).
[35] R. S. Ferreira and S. C. Ferreira, Critical behavior of the contact process on small-world networks, Eur. Phys. J. B **86**, 462 (2013).
[36] N. Friedman, S. Ito, B. A. W. Brinkman, M. Shimono, R. E. Lee DeVille, K. A. Dahmen, J. M. Beggs, and T. C. Butler, Universal critical dynamics in high resolution neuronal avalanche data, Phys. Rev. Lett. **108**, 208102 (2012).
[37] A. J. Fontenele, N. A. P. de Vasconcelos, T. Feliciano, L. A. A. Aguiar, C. Soares-Cunha, B. Coimbra, L. Dalla Porta, S. Ribeiro, A. J. Rodrigues, N. Sousa, P. V. Carelli, and M. Copelli, Criticality between cortical states, Phys. Rev. Lett. **122**, 208101 (2019).
[38] B. Mariani, G. Nicoletti, M. Bisio, M. Maschietto, S. Vassanelli, and S. Suweis, Disentangling the critical signatures of neural activity, Sci. Rep. **12**, 10770 (2022).
[39] M. Keeling, The implications of network structure for epidemic dynamics, Theor. Pop. Biol. **67**, 1 (2005).
[40] S. Valverde, S. Ohse, M. Turalska, B. J. West, and J. Garcia-Ojalvo, Structural determinants of criticality in biological networks, Front. Physiol. **6**, 6127 (2015).
[41] L. de Arcangelis and H. J. Herrmann, Self-organized criticality on small world networks, Physica A **308**, 545 (2002).
[42] D. B. Larremore, W. L. Shew, and J. G. Restrepo, Predicting criticality and dynamic range in complex networks: effects of topology, Phys. Rev. Lett. **106**, 058101 (2011).
[43] J. G. F. Campos, A. de Andrade Costa, M. Copelli, and O. Kinouchi, Correlations induced by depressing synapses in critically self-organized networks with quenched dynamics, Phys. Rev. E **95**, 042303 (2017).
[44] S. Yu, A. Klaus, H. Yang, and D. Plenz, Scale-invariant neuronal avalanche dynamics and the cut-off in size distributions, PloS One **9**, e99761 (2014).
[45] A. Levina and V. Priesemann, Subsampling scaling, Nat. Commun. **8**, 15140 (2017).
[46] S. Yu, T. L. Ribeiro, C. Meisel, S. Chou, A. Mitz, R. Saunders, and D. Plenz, Maintained avalanche dynamics during task-induced changes of neuronal activity in nonhuman primates, Elife **6**, e27119 (2017).
[47] L. J. Fosque, R. V. Williams-García, J. M. Beggs, and G. Ortiz, Evidence for quasicritical brain dynamics, Phys. Rev. Lett. **126**, 098101 (2021).
[48] D. J. Korchinski, J. G. Orlandi, S.-W. Son, and J. Davidsen, Criticality in Spreading Processes without Timescale Separation and the Critical Brain Hypothesis, Phys. Rev. X **11**, 021059 (2021).
[49] A. Das and A. Levina, Critical neuronal models with relaxed timescale separation, Phys. Rev. X **9**, 021062 (2019).
[50] L. D. Porta and M. Copelli, Modeling neuronal avalanches and long-range temporal correlations at the emergence of collective oscillations: Continuously varying exponents mimic M/EEG results, PLoS Comput. Biol. **15**, e1006924 (2019).
[51] V. Buendía, P. Villegas, R. Burioni, and M. A. Muñoz, Hybrid-type synchronization transitions: Where incipient oscillations, scale-free avalanches, and bistability live together, Phys. Rev. Res. **3**, 023224 (2021).
[52] https://github.com/LevinaLab/spatial-branching-network.
[53] A. Levina, V. Priesemann, and J. Zierenberg, Tackling the subsampling problem to infer collective properties from limited data, Nat. Rev. Phys. **4**, 770 (2022).
[54] Other models such as the contact process can be simulated more efficiently due to their asynchronous nature, and the fact that

active individuals do not interact anymore with other active individuals.

[55] C. Gardiner, *Stochastic Methods: a Handbook for the Natural and Social Sciences*, Springer Series in Synergetics (Springer, Berlin, Heidelberg, 2009).

[56] This is done by seeing that $a_2(p_s = 0) > 0$, $a_2(p_s = 1) = 0$, and then performing derivatives until it is clear that the function is monotonous.

[57] A. Clauset, C. Shalizi, and M. Newman, Power-law distributions in empirical data, SIAM Rev. **51**, 661 (2009).

[58] A. Klaus, S. Yu, and D. Plenz, Statistical analyses support power law distributions found in neuronal avalanches, PLoS ONE **6**, e19779 (2011).

[59] J. Touboul and A. Destexhe, Power-law statistics and universal scaling in the absence of criticality, Phys. Rev. E **95**, 012413 (2017).