Criticality in Spreading Processes without Timescale Separation and the Critical Brain Hypothesis

Daniel J. Korchinski[1,*], Javier G. Orlandi[1,†], Seung-Woo Son[1,2,‡] and Jörn Davidsen[1,3,§]

[1]*Department of Physics and Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada*
[2]*Department of Applied Physics, Center for Bionano Intelligence Education and Research,*
Hanyang University, Ansan, 15588, Republic of Korea
[3]*Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta T2N 4N1, Canada*

Spreading processes on networks are ubiquitous in both human-made and natural systems. Understanding their behavior is of broad interest: from the control of epidemics to understanding brain dynamics. While in some cases there exists a clear separation of timescales between the propagation of a single spreading cascade and the initiation of the next—such that spreading can be modeled as directed percolation or a branching process—there are also processes for which this is not the case, such as spiking cascades in neural networks. For a large class of relevant network topologies, we show here that in such a scenario the nature of the overall spreading fundamentally changes. This change manifests itself in a transition between different universality classes of *critical* behavior, which determines the onset and the properties of neural activity turning epileptic, for example. We present analytical results in the mean-field limit, giving the critical line along which scale-free behavior can be observed. The two limits of this critical line correspond to the universality classes of directed and undirected percolation, respectively. Outside these two limits, this duality manifests itself in the appearance of critical exponents from the universality classes of both directed and undirected percolation. We find that the transition between these exponents is governed by a competition between merging and propagation of activity, and we identify an appropriate scaling relationship for the transition point. Finally, we show that commonly used measures, such as the branching ratio and dynamic susceptibility, fail to establish criticality in the absence of timescale separation, calling for a reanalysis of criticality in the brain.

I. INTRODUCTION

Diseases spreading on human contact networks [1–3], worms and malware tunneling through computer networks [4], rumors shared on social networks [5], power failures cascading on electrical networks [6,7], and neuronal avalanches in the brain [8–15] are all spreading processes unfolding on a network. Such processes typically exhibit a directed-percolation (i.e., branching process) phase transition as the probability of spreading passes a critical threshold.

At this critical point, these systems exhibit shared scale-free statistics exhibiting characteristic power-law distributions of spreading cascades. Neuronal spreading cascades or avalanches with critical exponents similar to that of directed percolation have been observed experimentally in neural systems that differ in size by many orders of magnitude, from *in vitro* slices of a few hundred neurons, to whole-brain *in vivo* calcium imaging and functional magnetic resonant imaging (fMRI) [8,9,11,16,17]. These neuronal avalanches are at the core of the critical brain hypothesis and link together self-organizing principles in brain dynamics and connectivity with optimal information processing [13,15]. Equivalent power-law distributions have been found to emerge naturally in the course of training artificial neural networks, suggesting that they are a fundamental property of neural networks [18].

Although the mapping of neuronal avalanches to branching processes has proven quite successful, unavoidable discrepancies have appeared in recent years. There is an ongoing debate on whether these systems are really critical, quasicritical, subcritical, or a different definition altogether

*[*]djkorchi@ucalgary.ca
[†]javier.orlandigomez@riken.jp
[‡]sonswoo@hanyang.ac.kr
[§]davidsen@phas.ucalgary.ca

[19,20]. This debate stems at least in part from challenges identifying a suitable order parameter and determining whether observed avalanches are truly power-law distributed following the predicted mean-field values. Although an avalanche-size distribution $p(s) \sim s^{-\tau}$ with $\tau \approx 1.5$ (consistent with the mean-field exponents) has been widely reported in the literature [11,16,21], avalanches with exponents ranging from 1.2 to 2.5 have also appeared [10,14,22–24]. Whether the critical exponents even belong to directed percolation has also been challenged, with some proposing a oscillation-synchronization transition instead of a percolation transition [12,23–25]. These observations are complicated by the experimental limitations of sub-sampling and coarse-graining [26]. Further, experimental constraints make it challenging to distinguish subcritical avalanches from those truncated by finite-size effects.

Some of these issues, however, might be due to the often-overlooked fact that real systems rarely show proper time-scale separation; a classical branching process only allows for nodes to be excited when induced to do so by an antecedent node, and branching processes typically presuppose one "root" node. In other words, a branching process description assumes avalanches propagate and terminate on timescales much faster than the initiation of new avalanches. In neural systems, however, neurons can spontaneously activate due to internal stochastic dynamics, "minis" (the spontaneous release of neurotransmitter vesicles at the synaptic cleft), and external inputs, amongst other reasons. For smaller and intermediate systems, where multiple independent cascades are rare, the assumption of timescale separation is not particularly limiting. If the rate of spontaneous activity is not too high and all activity is aggregated into a single avalanche, then τ changes from the expected 1.5 to 1.25 [27]. However, in large neural systems, there are no global quiet periods with which to delimit neuronal avalanches as in smaller systems. One approach is to define avalanches as the periods during which activity exceeds a certain threshold, but it is not clear whether these thresholding methods identify a genuine critical point [12,25]. This uncertainty reflects the general challenge of defining criticality in strongly driven systems. However, some progress has been made. To disentangle neural avalanches in whole-brain fMRI and study their statistics, Tagliazucchi *et al.* used physical proximity (i.e., nearest-neighbor connections) to delimit avalanches, instead of binning all activity together [11]. The same method was necessary to assess criticality in whole-brain zebrafish data [17]. The approach of using network topology to identify avalanches makes sense, as information processing can only occur between connected elements of the network. To this end, neuronal avalanches have been generalized by "causal webs," which use the network structure to separate out independent avalanches [28].

It has remained an open question whether a genuine critical point with scale-free activity can exist alongside spontaneous activity, particularly as other markers of a directed-percolation transition (such as a unity branching ratio and the appearance of an active fraction) are all affected in different ways by spontaneous activity. We address the issue here using a minimal spreading process with a spontaneous activation rate, making it a discrete-time susceptible-infected-susceptible (ε-SIS) model in terms of the epidemic dynamics, and study it on a variety of complex networks.

Without spontaneous activations, there are no concurrent independent cascades of activity, and a directed-percolation phase transition is present. We show that the introduction of spontaneous activations means that the macroscopic markers used to identify the directed-percolation transition (i.e., the appearance of an active fraction, or a branching ratio of 1) no longer identify a phase transition. Nonetheless, by using the network structure to disentangle causally unrelated avalanches, we can define a phase transition even in the presence of spontaneous activations, with scaling relations between exponents and finite-size scaling. We perform an extensive study of the critical properties at this phase transition on a variety of network structures and show that the presence of any spontaneous activation changes the underlying universality class from that of directed percolation to that of undirected percolation, while preserving some features of directed percolation. To explain these results, we derive an analytical mean-field theory for branching processes with spontaneous activations and show that the appearance of undirected percolation exponents is a direct result of the merging of initially independent avalanches.

II. RESULTS

A. Model

To study the effect of spontaneous activations, here we consider a discrete-time ε-SIS process [29] on directed networks equipped with spontaneous activations (see Fig. 1). At each time step, a given node can be activated by means of a spontaneous activation with probability p or through an incoming link by an infected (activated) parent with probability q. More precisely, the probability that node i is activated at time $t + 1$ is given by

$$P(i, t + 1) = 1 - \bar{p}\bar{q}^{m(i,t)}, \qquad (1)$$

where $\bar{p} = 1 - p$ and $\bar{q} = 1 - q$ denote the complementary probabilities and $m(i, t)$ counts how many parents of node i were active at time t. In this model, activity in a node spreads to its neighbors in a single time step. Therefore, spontaneous activations occur on a timescale of p^{-1}, while the timescale of avalanches will depend on q, network structure, and p in a nontrivial way. This model can also be considered a type of Domany-Kinzel cellular automaton [30] or akin to some form of mixed percolation [31]. In our

model, nodes do not remain activated but recover and may be reactivated the following time step.

Large systems with spontaneous activations will have concurrent and possibly unrelated avalanches. We employ the network structure itself to identify independent avalanches, using the causal-webs approach described in Ref. [28]. To summarize the approach, we identify nodes with no active parents (i.e., no possible source of network-borne activation) as "roots" of newly initiated avalanches (see Fig. 1). Nodes with active parents inherit the avalanche labels of their parents. As avalanches overlap, they are merged together, so nodes only ever have one label. This merging reflects the fact that true causal information is often obscured in real systems or that contributions from both streams of activity are necessary for activation. This description of avalanches maps naturally to the clusters of traditional percolation.

The model has two limiting cases: (i) For $p = 0$, this model is a pure branching process with branching parameter q, which belongs to the directed-percolation universality class. This case corresponds to neural activity where avalanches are infrequent and a single leading neuron can be positively identified in each avalanche. (ii) Here, $q = 0$ corresponds to the ordinary percolation model on a directed network with probability p, which corresponds to neural activity that is entirely driven by external sources or spontaneous activation and does not spread on the network, such as in the retina. Both limits exhibit continuous phase transitions but fall into different universality classes and are characterized by power laws exhibiting different critical exponents. Our principle goals in this paper are to show that a phase transition exists for p and q, which are simultaneously nonzero, and to establish the universality class of this transition.

B. Numerical results

The most experimentally accessible indicator of criticality in systems with activity spreading is the size distribution of clusters, which shows a different critical exponent in directed and undirected percolation. For all p, at some critical $q_c(p)$, we observe a transition that defines a critical line. Below the critical point, with $q < q_c(p)$, avalanches are limited in size (see Fig. 2), while above the critical point, a permanent giant component affecting a nonzero fraction of the network appears. Note, however, that for all $p > 0$, there is a nonzero active fraction, even for $q < q_c(p)$. Therefore, this transition is not marked simply by the appearance of an active fraction, as is the case in a traditional absorbing phase transition. At the critical point, the exponential cutoff that characterizes the subcritical phase diverges, and the avalanche distribution is described asymptotically by a power law. The appearance of these power laws is used to identify the critical line in our simulations. By studying the critical exponent characterizing these power laws, we can identify the universality class of the critical line.

Because critical exponents depend on the network topology and dimension, we study avalanche distributions with extensive simulations on a variety of relevant network architectures, including small-world [32] [Figs. 3(a) and 3(b)], power-law networks [Figs. 3(c) and 3(d)], the hierarchical modular network [22] recently used as a brain connectome analogue [Fig. 3(f)], and the analytically tractable k-regular network [Fig. 3(e)]. Simulation details are explained in the Appendix A. Strikingly, for every level of spontaneous activity, there is a transition between two power-law exponents in the critical avalanche distribution.

We find that, in all cases, the first power-law exponent is consistent with the directed-percolation exponent for that

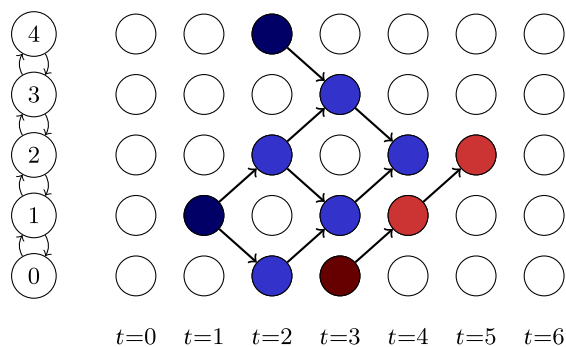$t=0$ $t=1$ $t=2$ $t=3$ $t=4$ $t=5$ $t=6$

FIG. 1. Example dynamics of the branching process with spontaneous activations. Multiple spontaneous activations are initiated on a simple linear bidirectional network (left). The dynamics here exhibit two independent avalanches, one with two roots (node 1 at time $t = 1$ and node 4 at time $t = 2$) and one with a single root (node 0 at time $t = 3$). With spontaneous activations, spatially distinct events can overlap in time (e.g., $t = 3$ and $t = 4$), and initially distinct cascades of activity can overlap to form larger avalanches.
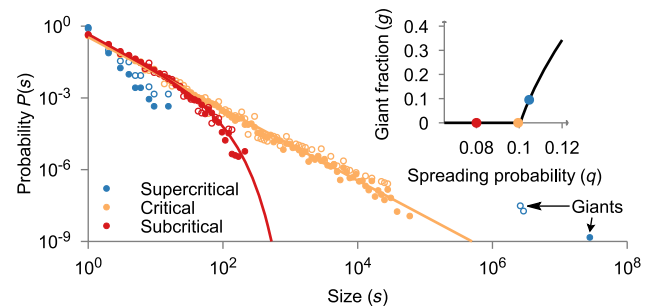

FIG. 2. Representative example of the directed-percolation phase transition. Avalanche statistics for low values of the spreading parameter q are an exponentially truncated power law $p(s) \sim s^{-3/2} \exp[-s/s_c]$. At criticality, with $q = q_c$, the cutoff diverges as $s_c \to \infty$, so $P(s) \sim s^{-3/2}$. For $q > q_c$, a fraction g of all active nodes are part of the giant component, an eternal avalanche (see inset). The absolute size of the giant avalanches therefore scales extensively with the network size $N = 10^4$ (empty circles), $N = 10^5$ (filled circles). Simulations are on a 10-regular directed network.
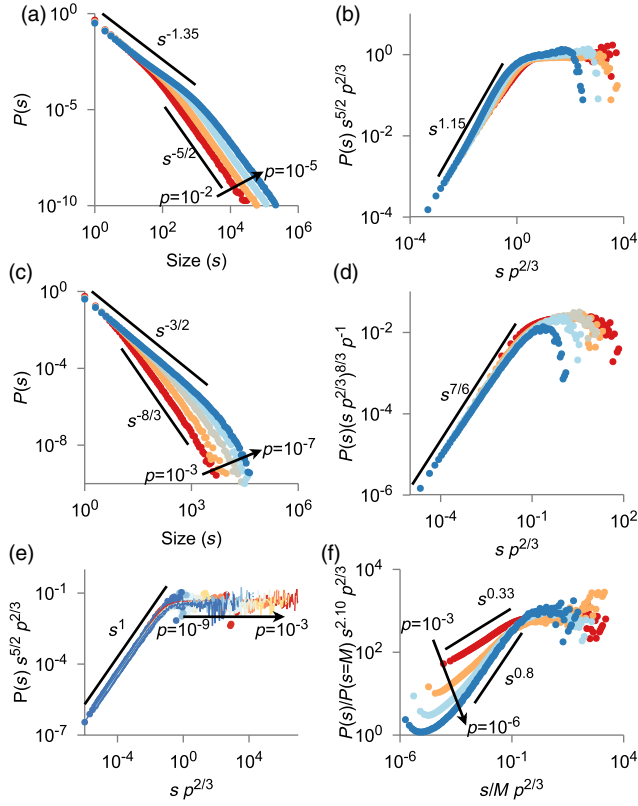
FIG. 3. Critical avalanche statistics on critical line for various network topologies. (a) Small-world networks with $N = 10^5$ and average degree $\langle k \rangle = 10$. (b) Same as in (a) but rescaled to produce a curve collapse. (c) Uncorrelated power-law in- and out-degree distributions, $p(k) \sim k^{-3.5}$, with $N = 10^7$ nodes. (d) Same as in (c) but rescaled to produce a curve collapse. (e) Rescaled avalanche distributions from 10-regular networks. Lines are simulations on infinite networks, while transparent circles are from finite simulations with $N = 10^7$; both are at $q = q_c(p)$, the analytically determined critical point. (f) Rescaled avalanche distributions from a hierarchical modular network, with $N = M \times 2^{15}$ nodes, on a 15-layer hierarchy with base module size $M = 10^2$. Black solid lines are guides for the eye. Unscaled panels (e) and (f) can be found in Figs. 16(a) and 16(b). The critical lines arising in each network topology are presented in Fig. 6 in Appendix A.

system, while the latter exponent is indistinguishable from the pure percolation exponent (see Table I). Yet, to the best of our knowledge, no directed-percolation avalanche exponent has been reported for directed small-world networks. To check the consistency of our findings for these networks, we can lower the density of long-range connections. Indeed, we find that the directed-percolation exponent for small-world networks tends toward the $(1 + 1)$-dimensional directed-percolation limit of approximately 1.108 expected of a circulant graph (see Fig. 12 in Appendix F). For the power-law network, the corresponding degree exponent is chosen such that the undirected-percolation exponent would change (from the 5/2 mean-field value to 8/3 as predicted in Ref. [33]) while leaving the directed-percolation exponent at 3/2 (as predicted in Ref. [34]).

In hierarchical modular networks [22], we observe non-scale-free behavior for avalanches below the base module size M ($s < M$), a p-dependent power law for intermediate-size avalanches ($M < s < M p^{-2/3}$), and finally, a single power law in the tail [cf. Fig. 3(f) in Appendix H]. The varying exponent for intermediate-size avalanches is consistent with reports of a Griffiths phase in modular networks with the SIS model [22,36,39], which belongs to the universality class of directed percolation. The largest avalanches are governed by an exponent of approximately 2.1, which matches with the undirected-percolation exponent for $q = 0$ [cf. Fig. 16(b)]. We hypothesize that this exponent is close to the pure two-dimensional percolation exponent because the hierarchical modular network has a backbone that is very nearly one dimensional, so the percolation process sees an effectively two-dimensional lattice upon the introduction of time. All critical-avalanche distributions exhibit a universal curve collapse for various p by rescaling the distribution by $p^{-2/3}$ [Figs. 3(b), 3(d)–3(f)], which indicates that for all $p > 0$, the critical point belongs to the same universality class for that network topology. In the supercritical regime, a giant component appears, just as in directed and undirected percolation. The probability that a randomly selected node is in the giant component is g, which exhibits a power-law scaling,

TABLE I. Calculated critical exponents for different network structures. In each column, we report both the theoretical value from percolation theory and the value determined in this work, either analytically (by application of generating functions) or from numerical simulations (where they are reported with decimal values). Errors in $1/\bar{\nu}$ indicate the range over which an acceptable curve collapse was obtained. Here, s_m is the crossover from the first, directed percolation, power law to the second. A number of additional critical exponents were determined for the k-regular network, which are summarized in Table III.

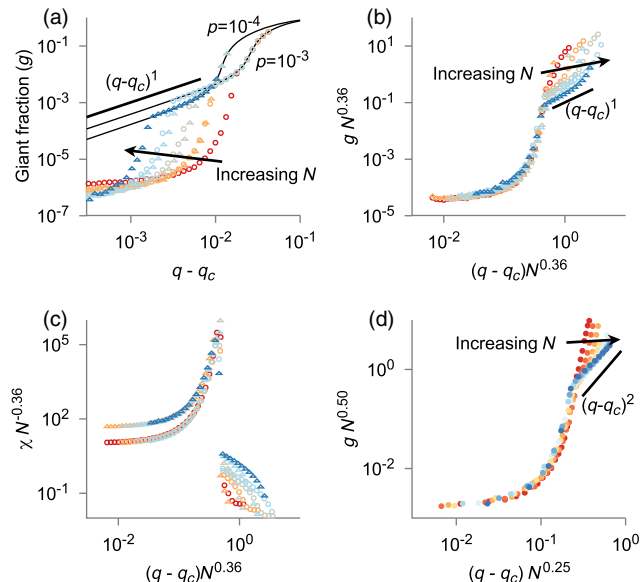| Exponent	τ_{DP}		τ		β		$1/\bar{\nu}$	
Quantity	$P(s) \sim s^{-\tau_{DP}}$ for $s < s_m$		$P(s) \sim s^{-\tau}$ for $s > s_m$		$g \sim (q - q_c)^\beta$		$q_c(N) - q_c \sim N^{-1/\bar{\nu}}$	
Small-world	$\cdots$	≈ 1.35	5/2 [35]	≈ 2.5	1 [35]	≈ 1.0	$\cdots$	0.35(5)
Power-law	3/2 [34]	≈ 1.5	8/3 [33]	≈ 2.67	2 [33]	≈ 2.0	1/5 [33]	0.25(5)
Hierarchical modular network	Varies [22,36]	Varies	$\cdots$	≈ 2.1	$\cdots$	≈ 0.8	$\cdots$	0.15(5)
k-regular network	3/2 [37]	3/2	5/2 [38]	≈ 2.5	1 [38]	1	1/3 [33]	0.36(2)

FIG. 4. Finite-size scaling effects on giant component size and susceptibility. (a) Giant components on 10-regular networks of varying sizes. Circles and triangles are $p = 10^{-3}$ ($N = 10^4$ to 10^6) and $p = 10^{-4}$ ($N = 10^{14/3}$ to $N = 10^{20/3}$), respectively, while solid lines are the analytical calculations for infinite lattices. (b) As in panel (a) but rescaled to produce a finite-size scaling curve collapse. (c) Same symbols as in panel (a), but studying the susceptibility $\chi = \langle s^2 \rangle_c$ for finite clusters, which shares the same finite-size scaling exponent. (d) Curve collapse for power-law networks [degree distribution $p(k) \sim k^{-3.5}$] with $p = 10^{-4}$ ($N = 10^{14/3}$ to N^7). The solid line is $g \sim (q - q_c)^\beta$ for $\beta = 2$, consistent with the theoretical prediction.

$g = (G/NT) \sim (q - q_c)^\beta$, where G denotes the size of the largest cluster, N the number of nodes in the system, and T the simulation duration. However, g exhibits a strong finite-size effect, with smaller systems having a larger effective critical point. Below the effective critical point, the largest cluster G does not scale with the simulation duration and is not percolating. For this reason, we resort to finite-size scaling to reveal the critical behavior of g on finite networks (Fig. 4). As $N \to \infty$, the effective critical point $q_c(N)$ tends towards the true critical point, with $(q_c(N) - q_c) \sim N^{-1/\bar{\nu}}$. The correctness of our finite-size scaling is confirmed by considering the finite cluster susceptibility $\chi \equiv \langle s^2 \rangle_c \sim (q_c - q)^\gamma$ (where the subscript c denotes an average over all clusters), which should obey the same scaling collapse with $\chi N^{-\gamma/\bar{\nu}}$ [Fig. 4(c)]. Mean-field theory predicts $1/\bar{\nu} = 1/3$, and for pure percolation on power-law networks [with degree distribution $p(k) \sim k^{-3.5}$], we expect that $1/\bar{\nu} = 1/5$ [33]. Here, we find that the best scaling collapse occurs near these values, with $(1/\bar{\nu}) \approx 0.36$ for 10-regular networks and $(1/\bar{\nu}) \approx 0.25$ for power-law networks.

Above the effective critical point, the giant component agrees with our analytical predictions for the infinite-size limit [Fig. 4(a)]. As expected, the giant components emerge with $\beta = 1$ for the mean-field case of random 10-regular networks [Figs. 4(a) and 4(b)]. As can be seen in Fig. 4(d), the giant component grows with $\beta = 2$ for the given power-law networks. Since $p(k) \sim k^{-3.5}$ and there are no correlations between the in-degree and out-degree, it is known that $\beta = 2$ for undirected percolation [33] and $\beta = 1$ for directed percolation [34], which implies that the emerging giant component in our system is in the universality class of undirected percolation.

C. Analytical results

In this section, we analytically establish that the universality class of the phase transition is lifted from directed to undirected percolation by the addition of spontaneous activations. To understand the transition between directed- and undirected-percolation exponents, we consider the analytically tractable k-regular network. Using the generating function formalism, we can explicitly derive scaling exponents related to the avalanche size and emergence of the giant component, as well as derive a critical line. By studying the sizes of singly rooted avalanches on this critical line, we can identify the size at which the merging of independent clusters of activity becomes the predominant mechanism for cluster growth and thereby explain the scaling collapse effected by $p^{2/3}$ observed in Figs. 3(c)–3(f). Additionally, we identify a diverging correlation time $\xi_\parallel \sim (q_c - q)^{-\nu_\parallel}$ at the critical point, with the undirected-percolation value of $\nu_\parallel = \nu = 1/2$ rather than the $\nu_\parallel = 1$ of directed percolation, reinforcing that this is an undirected-percolation transition.

The avalanche distribution of our model is akin to the cluster-size distribution of percolation and directed percolation; this distribution has been analytically determined on a variety of infinite random networks for both types of percolation by using probability generating functions (PGFs) [1,33,34,40]. The technique's key assumption is that there are no loops and that all nodes are equivalent, in the sense that their network properties are independent of the properties of their neighbors. Although this is only an approximation, this treelike approximation can still perform well in cases where loops are prevalent [41]. This assumption lets one write down a self-consistent equation for the PGF in terms of the number of connected neighbors where the cluster that each neighbor connects to is distributed according to the original PGF. Our approach, detailed in Appendix B, follows that same spirit, except that a system of two self-consistently coupled PGFs is required to describe the total cluster distribution. One PGF corresponds to the sizes of clusters reached from a direct descendant, while the other describes the cluster size reached when two independent cascades merge. On a treelike network, merging occurs when spontaneous activations meet and become a larger avalanche. Therefore, the

directed-percolation-like behavior is entirely contained within the first PGF, while the second PGF captures the effect of new spontaneous activations.

This pair of PGFs combines to define the PGF $H_0(x) = \sum_{s=1}^{\infty} P_n(s)x^s$, which corresponds to the avalanche-size distribution $P_n(s)$ obtained from sampling random active nodes (denoted with the subscript n). Derivatives of $H_0(x)$ evaluated at $x = 0$ directly yield $P_n(s)$. Since the giant component is the unique infinite-size avalanche, the probability that a random active node belongs to it is $1 - H_0(1)$, so the giant component is also determined by H_0. The average cluster size is just given by $\langle s \rangle_n = H_0'(1)$ and the susceptibility $\chi = \langle s^2 \rangle_c = H_0'(1)/\int_0^1 H_0(x)dx$, where the subscript c indicates sampling over clusters as opposed to active nodes. Analytically, we find

$$\chi \sim \langle s \rangle_n = \frac{1 - \frac{\sigma}{k}(\sigma - \sigma_m)}{(1 - \sigma)^2 - \frac{k-1}{k}\sigma\sigma_m}, \qquad (2)$$

where $\sigma(p, q)$ is the reproduction number or branching ratio, and $\sigma_m(p, q)$ is the merging number, corresponding to the number of cascades of activity leading to a randomly selected active node with at least one parent (see Appendix B 3 for details). For sufficiently small p and q, the giant component is zero and all clusters are finite, but as the critical line is approached, the susceptibility χ diverges as $\chi \sim |q_c - q|^{-1}$ (for fixed p) or $\chi \sim |p_c - p|^{-1}$ (for fixed q), as derived in Appendix B. This divergence defines the critical line and occurs when

$$0 = k(1 - \sigma)^2 - (k - 1)\sigma\sigma_m. \qquad (3)$$

Clearly, then, the critical line has $\sigma < 1$ for all $\sigma_m > 0$, meaning that giants can occur even before an average reproduction number of 1 is attained.

As shown in Fig. 5(a), $\sigma = 1$ on the critical line only in the $p \to 0$ and $q \to (1/k)$ limit of directed percolation for a k-ary tree [38]. At the directed-percolation critical point, the active fraction of nodes $\Phi = \langle \Phi(t) \rangle$ exhibits a divergence in its dynamic susceptibility χ_0, with $\chi_0 \equiv (\partial\Phi/\partial p)$ diverging as $\chi_0 \sim |(1/k) - q|^{-1}$. In the context of neural systems with mixed timescales and a fixed level of spontaneous activation, the maximum of this dynamic susceptibility defines a "Widom" line (cf. Fig. 10 and associated text in Appendix D) and has been proposed as a quasicritical line [42]. Although all three of these measures identify the directed-percolation critical point $p = 0$ and $q = (1/k)$, they disagree as soon as spontaneous activation is introduced ($p \neq 0$) and exhibit distinct scaling (cf. Fig. 11 in Appendix E). In the $p \ll 1$ limit, the Widom line scales as $p \sim [(1/k) - q]$, the $\sigma = 1$ line scales as $p \sim [(1/k) - q]^2$, and the critical line scales as
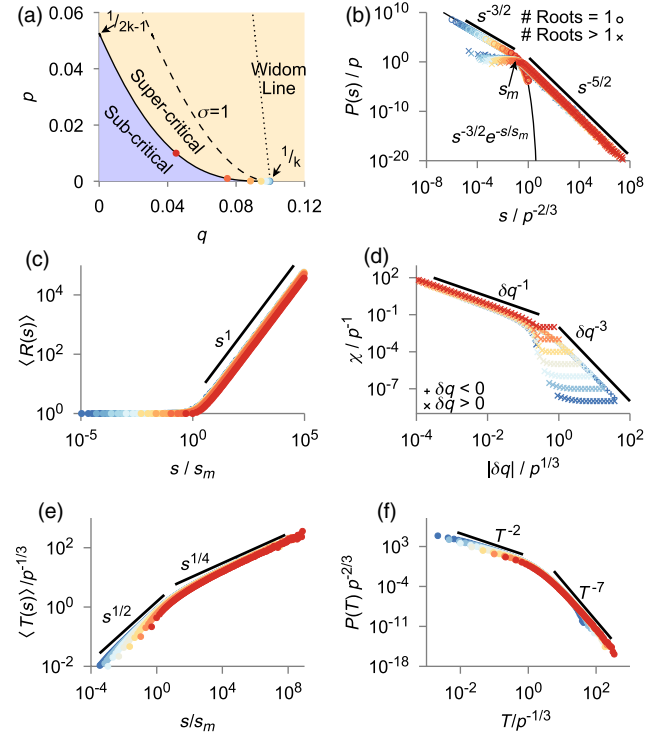

FIG. 5. Power-law transitions governed by merging. (a) Phase diagram for the k-regular network, with $k = 10$. Points on the critical line correspond to the $[p, q_c(p)]$ in the other panels of this figure, with the p ranging from 10^{-2} to 10^{-9}. (b) Rescaled avalanche-size distribution for various p simulated on an infinite 10-regular network, partitioned into those avalanches with a single initiation site (empty circles) and those with multiple initiation sites (crosses). The theoretical distribution of mergeless avalanches is indicated with the solid line [cf. Eq. (C2)]. (c) Rescaled average number of roots R for avalanches of a given size for simulations on an infinite 10-regular network. (d) Rescaled susceptibility near the critical point, where $\delta q = q_c - q$, calculated by the generating function H_0. Subcritical values $\delta q < 0$ are shown with empty circles and exhibit two power laws, while supercritical values $\delta q > 0$ show only one. (e) Average avalanche duration for simulations of a given size collapse onto a single curve. (f) Avalanche duration distribution collapsing onto a single curve with two power laws. Unscaled data for panels (c)–(f) are found in Fig. 17.

$$\left(\frac{1}{k} - q_c\right)^3 \approx \frac{(k-1)^2(2k-1)}{k^5}p_c. \qquad (4)$$

As for the $q = 0$ endpoint to the critical line, χ diverges when $p = [1/(2k - 1)]$ and $q = 0$, the pure percolation critical point for the Bethe lattice of coordination number $2k$. Hence, the critical line contains members belonging to two distinct universality classes.

To understand the appearance of the $p^{-2/3}$ scaling of the transition point shown in Fig. 3, we can consider the distribution of avalanches with only one root. These avalanches are described by a branching process, on a

k-ary tree, with a branching probability $P_{d1} = \bar{\Phi}^{k-1}(1 - \bar{p}\,\bar{q})$ corresponding to the probability that a daughter branch activates with exactly one parent. The probability distribution for the size of the singly rooted avalanche is $P_{\text{mergeless}}(s) \sim s^{-3/2} \exp[-s/s_m]$, where $s_m = -1/\ln\{kP_{d1}[\overline{P_d}/(1 - 1/k)]^{k-1}\}$ [see Eq. (C3)] denotes the characteristic scale above which avalanches merge and $\overline{P_d}$ is the probability a site does not activate, despite having an active parent. Hence, we expect that the exponent $s^{-3/2}$ should be exponentially suppressed at s_m. In the limit of $p \to 0$ on the critical line [see Eq. (C5)], s_m scales as

$$s_m \approx \frac{2(k-1)}{3k^3}\left(\frac{1}{k} - q\right)^{-2}. \qquad (5)$$

Combining Eqs. (4) and (5) shows that the characteristic size before merging scales as $s_m \sim p^{-2/3}$ on the critical line.

Simulations confirm that the smallest avalanches typically only have one root [Fig. 5(b)], while the largest avalanches have a number of roots that scale with the avalanche size [Fig. 5(c)]. Thus, there are two competing processes at play in these avalanches, both the propagation of the avalanche, which belongs in the directed-percolation universality class, and the merging of initially independent events, which falls into the percolation universality class. This competition explains the appearance of the two power laws and the associated curve collapse in Fig. 3. The first power law is governed by the spreading of activity from a single initiation site, while the second power law is governed by the merging of activity springing from multiple sites. This $-\frac{2}{3}$ scaling is a good approximation for random graphs that are close to the mean-field limit. In Appendix F, we consider small-world networks with a low shortcut density. These networks are locally one dimensional and, as a consequence, exhibit a different scaling, $s_m \sim p^{-0.75}$ (cf. Fig. 12) due to the directed-percolation phase being $(1 + 1)$-dimensional instead of the mean-field limit.

This transition between the directed and undirected exponents also manifests itself in the approach to the critical point. For instance, for $q - q_c = \delta q < 0$, the susceptibility $\chi = \langle s^2 \rangle_c$ can be approximated using $P(s) \sim s^{-\tau_{\text{DP}}}\mathcal{F}_1(s/s_m) + \Theta(s_\xi - s_m)s^{-\tau}\mathcal{F}_2(s/s_\xi)$, where the $\mathcal{F}_{1/2}$ represent universal scaling functions, Θ is the Heaviside step function, and $s_\xi \sim |\delta q|^{-1/\sigma}$ is the size cutoff of the pure-percolation tail. Then,

$$\chi \approx \int_1^\infty s^2 s^{-\tau_{\text{DP}}}\mathcal{F}_1\left(\frac{s}{s_m}\right)ds + \Theta(s_\xi - s_m)\int_1^\infty s^2 s^{-\tau}\mathcal{F}_2\left(\frac{s}{s_\xi}\right)ds,$$

so $\chi \approx s_m^{3-\tau_{\text{DP}}} + s_\xi^{3-\tau}$; thus, using that $s_m \sim (1/k - q)^{-1/\sigma_{\text{DP}}} \sim (p^{1/3} + \delta q)^{-1/\sigma_{\text{DP}}}$ from Eqs. (4) and (C4), we obtain (up to arbitrary multiplicative constants C_1, C_2)

$$\chi = C_1 p^{-(3-\tau_{\text{DP}}/3\sigma_{\text{DP}})}(1 + \delta q/\sqrt[3]{p})^{-(3-\tau_{\text{DP}}/\sigma_{\text{DP}})}$$
$$+ C_2 \Theta(s_\xi - s_m)\delta q^{-(3-\tau/\sigma)},$$

which suggests that we see a transition between exponents when $\delta q \approx \sqrt[3]{p}$, precisely as observed in Fig. 5(d). Furthermore, since $\chi \sim \delta q^\gamma$ defines γ, we have arrived at the usual scaling relation $\gamma = (3 - \tau/\sigma)$, which holds for both the directed ($\gamma_{\text{DP}} = 3$, $\sigma_{\text{DP}} = \frac{1}{2}$, and $\tau_{\text{DP}} = \frac{3}{2}$) and undirected ($\gamma = 1$, $\sigma = \frac{1}{2}$, and $\tau = \frac{5}{2}$) percolation regimes.

A transition from directed-percolation exponents also appears in the dynamical exponents relating the size of avalanches to their duration [cf. Fig. 5(e)], where the exponent transitions from $s \sim \langle T \rangle^{\sigma\nu z = \frac{1}{2}}$ to a power law consistent with $s \sim T^{\frac{1}{4}}$. The onset of this transition again occurs with avalanches of size $s_m \sim p^{-\frac{2}{3}}$, which defines a characteristic time to merging, $T_m \sim \sqrt{s_m} \sim p^{-\frac{1}{3}}$. The scaling of this characteristic time captures an exponent transition in the distribution of the avalanche durations [cf. Fig. 3(f)], with $P(T) \sim T^\alpha$, with $\alpha_{\text{DP}} = 2$ and a new asymptotic $\alpha \approx 7.0$. Intriguingly, the directed-percolation scaling relation $(\tau - 1/\alpha - 1) = \sigma\nu z$ is satisfied even in the merging regime, assuming $\alpha = 7$, $\tau = 5/2$, and $\sigma\nu z = \frac{1}{4}$.

The existence of robust scaling relations and of curve collapses [Figs. 5(b)–5(f)] that appear universal indicates that the critical line (for $p > 0$) belongs to a single universality class. Since this includes the point $p = (1/2k - 1)$ and $q = 0$, which we know is exactly undirected percolation, it suggests that the entire critical line (save for $p = 0$, $q = 1/k$) belongs to the universality class of undirected percolation.

We can further strengthen the argument that the critical line is an undirected-percolation transition by studying the correlation lengths of the system. Undirected percolation exhibits a single isotropic diverging correlation length $\xi \sim |\delta q|^{-\nu}$, while directed percolation exhibits two diverging correlation lengths, $\xi_\perp \sim |\delta q|^{-\nu_\perp}$ and $\xi_\parallel \sim |\delta q|^{-\nu_\parallel}$, corresponding to spatial and temporal correlation lengths, respectively. We probe these correlation lengths and related ones using the two-point connectedness function, $\gamma(i, t_i, j, t_j)$, which measures the probability that node i at time t_i and node j at time t_j belong to the same cluster over the ensemble average. If the shortest path connecting nodes i and j has length d_{ij}, then we expect that the average connectedness function should decay with d_{ij}. This decay can be seen by studying the exponential decay of the average connectedness function, $c(d, t) = \langle \sum_{j \text{ for } d_{ij}=d} \gamma(i, t_i, j, t_j)\rangle_{t=t_j - t_i, i \text{ active}}$, which measures the decay of active sites away from an active node.

We show the existence of a diverging spatial correlation length by analytically showing that $c(2d, 0) \sim \exp[-d/\ell_\perp]$ and that $\ell_\perp \sim \delta q^{-1}$. This length is given by $\ell_\perp = -1/\ln[\sigma^2(1 + \sqrt{\Upsilon})^2]$ (see Appendix G), which implies

that the perpendicular correlation length diverges when $(1 - \sigma)^2 = \sigma^2 \Upsilon$. This independently reproduces the critical line on which $\langle s \rangle_n$ diverges. We stress that $\ell_\perp$, though it is a diverging spatial length scale that implies the divergence of $\xi_\perp$, is conceptually distinct from $\xi_\perp$. In infinite-dimensional systems such as ours, the latter is defined based on the characteristic cluster size and the associated "fractal" dimension [38].

For the temporal correlation length, one candidate is the direct-descendant correlation time ξ_t defined by $c(t, t) \sim \sigma^t \sim \exp[-t/\xi_t]$, which diverges when $\sigma = 1$. Yet, this definition neglects the contribution of merging. To capture the effect of merging, we define $\xi_\parallel \sim \sum_d c(d, t) = c(t)$, which counts the average number of connected nodes that are t time steps from an active node [30]. We find (see Fig. 15) that for a cutoff with $\delta q > p^{1/3}$, $\xi_\parallel \sim \delta q^{-1.0}$ (consistent with directed percolation), while for $\delta q < p^{1/3}$, $\xi_\parallel \sim \delta q^{-0.5}$ (consistent with isotropic percolation). This result precludes the directed-percolation transition except at $p = 0$, and it is consistent with the scaling of χ.

In summary, the critical line is an undirected (as opposed to directed) percolation transition except at the singular point $p = 0$. This is supported by avalanche distribution exponents, exponents of the order parameter g, undirected-percolation scaling relations, and the scaling of the temporal correlation length. Many critical exponents of the directed percolation remain observable on small scales, such as in the beginning of the avalanche-size distribution or in the susceptibility χ. These exponents then shift to the undirected exponents when the merging of initially independent avalanches becomes prevalent. Meanwhile, other measures of criticality that hold for directed percolation, such as the divergence of the dynamical susceptibility χ_0 and a reproduction number of 1, no longer capture critical behavior. Instead, they predict phase curves that agree only in the $p = 0$ limit and scale with different power laws near the directed-percolation limit. Specifically, $(1/k) - q \sim p^a$, with $a = 1$ for the Widom line, $a = 2$ for the $\sigma = 1$ line, and $a = 3$ for the critical line (cf. Fig. 11 in Appendix E). Thus, the directed-percolation transition is not robust with respect to the introduction of spontaneous activation—any level of exogenous driving will introduce independent outbreaks, which, on the largest scales, will begin to merge.

III. DISCUSSION

A. General model observations

We have described a two-parameter spreading process that includes spontaneous activations and exhibits a phase line along which the critical exponents and behavior of both directed and undirected percolation appear. When there is no spontaneous activation, the model exhibits a directed-percolation transition, marked by a divergence in the dynamical susceptibility, a reproduction number of 1, power-law distributed outbreak sizes, and the appearance of a giant component. However, the introduction of spontaneous activation means that the dynamical susceptibility no longer diverges and that the reproduction number is shifted. Nonetheless, by considering the cluster-size distribution and statistics related to the cluster size, a critical line—exhibiting universal curve collapses, diverging correlation lengths, and finite-size scaling—can be defined, which means that, even in the presence of spontaneous activity, a genuine phase transition exists.

The introduction of spontaneous activity destroys the transition to the absorbing state and shifts the phase line into the universality class of undirected percolation. This result is perhaps surprising because the undirected-percolation limit $q = 0$ obeys detailed balance and is an equilibrium phase transition, while for the remainder of the critical line with $q > 0$, detailed balance is not respected and the system is a nonequilibrium one. We showed numerically, on a variety of relevant network topologies, that in the largest clusters, merging becomes increasingly prevalent, causing the undirected-percolation exponents to dominate. The dominance of the undirected-percolation exponents suggests that under renormalization group flow, q is an irrelevant parameter, meaning the time asymmetry introduced by q is destroyed and the phase transition is *effectively* an equilibrium one.

As such, whether the time-directed *spreading* can be called critical can be debated. Although the temporal correlation length $\xi_\parallel$ diverges (with the percolation exponent $\nu = 1/2$), the direct-descendant correlation time $\xi_t \sim -1/\log(\sigma)$ is finite along the critical line when $p > 0$. This implies that, unlike at the directed-percolation $p = 0$ critical point, an active site can only propagate information over a finite time interval.

Although all the networks we considered were nominally directed, we expect that our results survive on undirected networks, as our small-world networks are comprised predominately of bidirectional connections.

B. Critical brain hypothesis

Our results have repercussions for the critical brain hypothesis. Although the hypothesis itself is not new [43,44], it gained traction with the seminal work on neuronal avalanches [16,21] and with the development of large-scale brain recording techniques. It is still a highly debated topic [19,20], and recent work has focused mostly on (i) the appropriate definition of an order parameter and its tuning and (ii) whether neural activity distributions show critical power-law statistics.

Regarding point (i), the main objective is to find a plausible mechanism by which the brain is able to tune its own activity to a critical point; ongoing research focuses

on self-organized criticality [45,46], excitatory-inhibitory activity balance [47], up and down states [48,49], adaptive mechanisms [50], and learning [51], amongst other topics. Many of these concepts are deeply related, and all of them might play a role. Regarding point (ii), early work focused on whether the measured statistics followed a real power law or just an approximation, and whether the activity was subcritical or supercritical instead [52]. However, this is a challenging issue to solve experimentally because of the role of finite-size and subsampling effects [53–55] and because of the real lack of separation of timescales as we report here. The initial studies on neuronal avalanches reported an exponent close to $\tau \approx 1.5$ for the size distributions, consistent with a mean-field branching process [9,16], but experiments on a variety of neural systems have reported a range of exponents, usually in the range 1.2–2.5. While this discrepancy might be partly explained by the technical challenges in probing the tails of power-law distributions or attributed to variation in network topology between studies and heterogeneous dynamical properties [14], there remain reasons to think spontaneous activation has a significant bearing on the critical brain hypothesis. For instance, critical *in vivo* neuronal avalanches have only been reported for waking states, when animals presumably are stimulated by sensory input [48,52,56]. In light of our work, this might indicate that criticality is only attained with the help of an external drive.

Our work addresses both points (i) and (ii). For point (i), we show that a susceptibility χ based on causal webs accurately identifies the critical line, with the excellent finite-size scaling one expects from a true phase transition. This approach offers a well-defined observable that is coherent in the presence of spontaneous activity, unlike the branching ratio or a global measure like the dynamic susceptibility. As for point (ii), we show that power laws (with exponents that vary based on network structure) are present at the critical point for any level of spontaneous activation. While power-law statistics can also appear in noncritical systems with simpler dynamics as a recent critique showed [57], the presence of scaling relations between the critical exponents [9] is only true in pure critical systems. We demonstrate such scaling relations in our model with coexisting power-law regimes, showing, for the first time, that neural networks with spontaneous activity can still be genuinely critical. Additionally, our prediction of two power laws may help to explain the variety of exponents fitted to power laws in the literature.

Much of the literature definitions and requirements for criticality cannot be satisfied in the presence of spontaneous activity since timescale separation is barely satisfied for any realistic activity rate. Only recently have other works started to focus on the role of spontaneous activity for the critical brain hypothesis [27,28,42,58–62]. Similar to the causal webs we use here, the introduction of "causal avalanches" allows for independent cascades to overlap in time when in the active phase, but it explicitly prevents the merging of clusters [59]. Because causal avalanches cannot overlap, they compete for activations, and their statistics are governed by neutral theory. This approach produces directed-percolation avalanche exponents in the limit of vanishing input, but nonuniversal statistics appear as the external input is increased [59].

Several approaches have been taken to allow only one avalanche at a given time. In the limit where spontaneous activations only contribute to an already existing avalanche, spontaneous activations can change avalanche exponents from directed percolation while leaving them otherwise scale-free [27]. Others have attempted to recover the original definition of criticality and power-law statistics by reducing the temporal bin size for the avalanche definition [14]; but it is still not clear whether that approach really recovers the same underlying dynamics, and it might only hold if the system is exactly at the critical point and for intermediate-size systems. Others have attempted to restore the classical definition by introducing a "threshold" to only count avalanches as activity above the steady-state active fraction Φ; however, this approach relies on an *ad hoc* threshold and can alter measured critical exponents [60]. Finally, even with self-tuning mechanisms tying the system close to the $\sigma = 1$ critical line for several decades in external driving strength, the requisite of nonoverlapping avalanches means scale-free avalanches are observed only in the limit of vanishing external drive [63].

It is clear that in the thermodynamic limit, with a fixed spontaneous activation rate, there will always be unrelated avalanches occurring, so avalanches defined by global observables (such as fluctuations in Φ or those delimited by global quiescence) are not well defined. As shown in the present work, in the absence of timescale separation, it is essential to know the network structure to resolve the underlying dynamics. There is currently no way to recover the correct exponents without access to the network structure, but approaches that first try to infer the structure from the dynamics appear to be promising [28].

Another question of interest in the context of the critical brain hypothesis is the effect of (weakly) correlated external inputs. In our minimal model, the spontaneous activity mimics a random Poisson process with no quenched disorder in activation rate, and the effect of any initially correlated input would die away. However, the rate of spontaneous activations in real neurons due to minis likely depends on the number of input synapses. Moreover, brains typically operate on correlated inputs and pass information between brain modules also by means of correlated inputs [64], which is rather far from the random inputs we consider in our model. Yet, some of our findings are expected to hold under weakly correlated inputs based on our simulations on hierarchical modular networks. In

these networks, the connections between modules introduce correlated inputs to the modules. The effect of these correlations seems to vanish, and avalanches appear scale-free for sizes larger than the typical module size [Fig. 3(f)]. The robustness of criticality to continuous and more strongly correlated external inputs should be studied further. Another key point to tackle in the future in relation to the critical brain hypothesis in the presence of spontaneous activity has to do with information transmission. Can it still be optimal in this regime? Spontaneous activity can indeed enhance information transmission from a sensory system [65] or shift an inherently subcritical system closer to the critical point (see Fig. 5). However, a read-out of this activity might require an error-correction code.

C. Other applications

The generality of our model makes it applicable to other systems where the timescales of spontaneous activation and propagation of activity are comparable, e.g., rumor spreading on social networks [5] or the distribution and propagation of computer viruses [4]. It remains to be seen how our findings translate to self-organized systems and, in particular, to those that are known to exhibit a self-organized critical (SOC) regime under timescale separation [66–68]. Previous studies have shown that a sufficiently high driving rate can induce a transition from avalanche dynamics to continuous flow in SOC systems [69].

Our results also apply to epidemiology and disease-spreading processes. A small number of studies have relaxed the patient-zero assumption of the typical spreading process by including multiple initial spreaders [70–76]. A limited few reflect a disease reservoir that can cause new outbreaks even as old ones spread [29,77,78]. This might be an appropriate description for diseases like Zika virus, which can spread via human sexual networks but also "off-network" via mosquito [79]. Alternatively, it might describe the population of a nation under lockdown, with occasional exogenous disease input. To the best of our knowledge, the question of the universality class in such situations has not yet been answered in the literature. Within the context of zoonotic diseases, a multilayer network approach that incorporates human-animal interactions directly [80] or includes cooperative diseases [81] might open the door for novel dynamics in the absence of timescale separation. It would also be interesting to see which metrics, such as k-shell decomposition [82] or a local analysis [83], identify significant spreaders in our model and if network interventions, such as those connected to explosive percolation [84–86] and others [87], could be used to stymie or promote epidemics in the presence of spontaneous activation. These remain exciting challenges for the future.

IV. CONCLUSIONS

Spreading processes on networks frequently appear in natural and human systems. The inclusion of spontaneous activity changes the phase transition in these systems from directed percolation to isotropic percolation because previously independent streams of activity can merge together. These universality classes have differing critical exponents, meaning that activity clusters will show different growth profiles near the critical point. Our findings have several implications for the critical brain hypothesis. Global quantities—such as the active fraction and its susceptibility, the branching ratio, or avalanches defined by global periods of quiescence—do not capture critical behavior when spontaneous activity is considered. As such, criticality in the brain should be reassessed using measures that tolerate spontaneous activity, which requires that the use of network structure (e.g., tractography) be paired with dynamics measurement (e.g., fMRI). Proximity to criticality should be assessed using an order parameter based on causal webs, such as susceptibility χ. Other measures of criticality, like the branching ratio, might lead to critical behavior being interpreted as subcritical. If the brain as a whole is critical, then the largest avalanches will have isotropic, rather than directed, percolation critical exponents as merging becomes the dominant growth mechanism.

APPENDIX A: METHODS

1. Network generation

For the finite networks, we generate finite directed k-regular networks via the configuration model, shuffling connections to avoid self-links and multilinks. To generate power-law networks, we employ a variation of the configuration model described in Ref. [88], with degree distributions $p^{\text{in/out}}(k) = k^{-3.5}$ within a domain $k \in [5, \ldots, 1000]$, with rejection parameters ($\kappa = 0.5$, $\delta = 0.05$ [89]). We generate the small-world networks using a directed network generalization of the Watts-Strogatz

model [32], using rewire probability 10^{-2} and with average degree 10 (a rewire probability of 10^{-3} is also considered in Appendix F). We generate the hierarchical modular networks described as "HMN-2" in Ref. [22] as a backbone for our modular networks. Within each base module with n_c connections in the module backbone, we place $10^2 + 4n_c$ nodes so that each intermodular connection is supported by four nodes. The first $M = 10^2$ nodes we draw from the out-degree distribution $p(k) \sim e^{-(k-10)^2/(2\times0.5^2)}$ and connect to other uniformly drawn nodes in the same module. For the next $4n_c$ nodes, we draw from the same out-degree distribution but connect to the first 10^2 nodes in the other modules, according to the module backbone wiring.

2. Simulation of model on finite networks

Networks are initiated with no active nodes. Each time step, the number of nodes that will activate is drawn from the binomial distribution, with activation probability p. That number of nodes is randomly selected with uniform weighting, redrawing duplicates. Nodes that activate spontaneously and had no active parents in the preceding time step initiate a new cluster. Then, all nodes that had active parents in the preceding time step that were not already activated spontaneously are checked for activation. Each node with m active parents in the previous time step is activated with probability $1 - \bar{q}^m$. Nodes inherit the cluster label of their parents. If a node would inherit more than one cluster label, then those clusters are merged into a single cluster by relabeling all nodes belonging to the smaller cluster with the label of the larger cluster. Clusters that are found to have no active nodes in a given time step are terminated, and their size, duration, and number of roots are recorded.

3. Simulations on infinite k-regular networks

For the infinite networks, we begin at a randomly selected active node. We then check its immediate neighbors to see whether they are part of the same cluster. For those neighbors that are included, we then check their neighbors for inclusion. We can perform this process such that we need only count the number of unexplored neighbors, of which there are two types: (I) daughters that have not been checked for inclusion and (II) parents that are known to be included but whose neighbors have not been checked. If we are beginning from a root node, there are initially k unchecked daughter branches (type I). If we are beginning from a randomly active node, we begin with one type-I neighbor and one type-II neighbor. The algorithm proceeds to check each unevaluated connection (of type I or type II), possibly adding more as it goes, until none remain or the cluster exceeds a given size (typically 10^{10}). Each type of connection is added as follows:

(i) Type I: We check each type-I neighbor, by assuming it has m_d other active parents [drawn from a binomial distribution $P(m_d) = \binom{k-1}{m_d} \Phi^{m_d} \bar{\Phi}^{k-1-m_d}$ of $k-1$ other parents, activated with probability Φ, the active fraction, given by Eq. (B11)], and include each type-I neighbor with probability $1 - \bar{p}\bar{q}^{m_d}$. If it is included, then we add k type-I connections from this daughter and m_d type-II connections.

(ii) Type II: Each type-II neighbor is included with probability 1. It adds $k-1$ additional type-I connections and m_p type-II parents, with m_p drawn from the distribution in Eq. (A1):

$$p(m_p) = \frac{\binom{k}{m_p} \Phi^{m_p} \bar{\Phi}^{k-m_p}}{\Phi} (1 - \bar{p}\bar{q}^{m_p}). \qquad (A1)$$

The probabilities of adding a daughter or parent are as derived in Appendix B. For the purposes of measuring the two-point connectedness function, the above algorithm can be easily extended to also include the number of time steps by simply tracking how many times each active front has followed a daughter branch or a parent branch.

4. Critical-point determination

Accurate determination of the critical point is necessary to effect accurate finite-size scaling. In the case of the k-regular network, the critical point can be determined analytically. However, for the power-law, small-world, and hierarchical modular networks, determination of the critical point can be done in two ways. The naive approach is to simply tune q for fixed N and p until power laws appear in the avalanche distribution. However, this approach is prone to finite-size effects: For fixed p and N, the largest power laws in the $P(s)$ distribution will appear at the pseudocritical point, corresponding to a larger $q_c(N)$ value than the
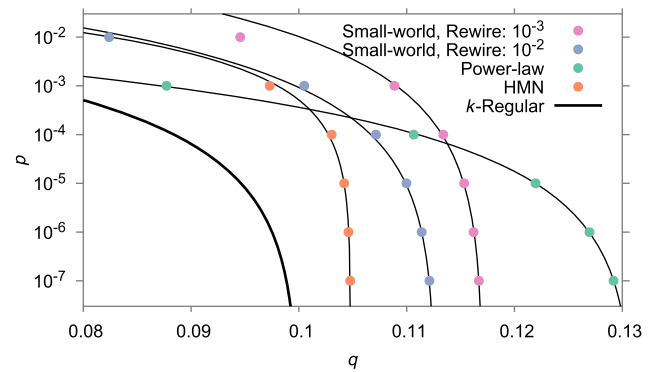

FIG. 6. Numerically determined critical lines for various networks. Points correspond to the avalanche simulations plotted in Fig. 3 in the main article, except for the small-world network with a rewire of 10^{-3}, which is only studied in Appendix F. Solid lines are approximate fits of the form $p(q) = c(q_{c,\mathrm{DP}} - q)^a$ for constant c, a, and $q_{c,\mathrm{DP}}$, whose values are summarized in Table II. The k-regular phase curve is exact.

TABLE II.   Fit parameters for the critical lines of Fig. 6. Fits are power-law fits of the form $p = c(q_{c,\text{DP}} - q)^a$. Entries for the k-regular network correspond to the low-p approximation given in Eq. (4).

Network	Fit parameters		
	c	a	$q_{c,\text{DP}}$
Small-world, rewire $= 10^{-2}$	6.00	2.97	0.112665
Small-world, rewire $= 10^{-3}$	7.99	3.09	0.117104
Hierarchical modular network	2.78	1.94	0.10482
Power-law network	2.29	2.94	0.131118
k-regular network	$k^5/(2k-1)(k-1)^2$	3	$1/k$

true $q_c(N \to \infty)$. Finite-size effects similarly cripple approaches based on just the appearance of the giant component or the diverging susceptibility.

Instead, we employ the method of D'Souza $et\ al.$ [90] and consider the quantity $B = g(\langle s^2 \rangle_c / \langle s \rangle_c^2)$, which was shown to have no finite-size dependence at the critical point. Therefore, for fixed p, the critical point q_c can be found as the intersection point of the $B(q)$ curves for different values of N [see Fig. 12(a)]. This procedure enables the numerical determination of the p, q critical lines for the networks discussed in the main text (cf. Fig. 6 and Table II).

APPENDIX B: GENERATING FUNCTIONS

Before deriving the generating functions associated with the average cluster size, we begin with a brief review of probability generating functions (PGFs). For a discrete random variable X drawn from the probability mass function $p(x)$, the probability generating function can be defined as

$$g_X(z) = E(z^X) = \sum_{x=0}^{\infty} p(x)z^x.$$

Here, g_X generates the probability $p(x)$ in the sense that $g_X(0) = p(0)$, and the nth derivative yields $(1/n!)g_X^{(n)}(0) = p(n)$. The probability generating function can be used to obtain the moments of X, as $\langle X \rangle = g_X'(1)$, $\langle X(X-1) \rangle = g_X''(1)$, and so on. The final property of probability generating functions we use is perhaps its most useful: When a family of independent and identically distributed variables $\{X_1, X_2, \ldots, X_N\}$ generated by $g_X(z)$ are summed, $Y = \sum_i^N X_i$, with N also being a random variable generated by $g_N(z)$, then $g_Y(z) = E(z^Y) = E(z^{NX}) = \sum_{n=1}^{\infty} p(n = N)(E(z^X))^n = g_N(g_X(z))$. Although this may seem esoteric, it means that the sum of a collection of some

random number of random variables can be concisely expressed using generating functions.

We derive the PGFs corresponding to the cluster-size distribution beginning from randomly selected active sites on a random network. If there are no loops in the network (the treelike approximation), we can express the cluster size s starting from a random active site as

$$s = 1 + \sum_{l=1}^{n_d} s_{d,l} + \sum_{m=1}^{n_p} s_{p,m}, \tag{B1}$$

where n_d is the number of active daughters of the initial site, n_p is the number of active parents of the initial site, $s_{d,l}$ is the size of the cluster reached from the lth active daughter, and $s_{p,m}$ is the size of the cluster reached from the mth active parent. Thus, the PGF for the total cluster size s is

$$H_0(x) = xA_o(H_p(x))B_i(H_d(x)), \tag{B2}$$

for the PGFs A_o (generating n_d), B_i (generating n_p), $H_p(x)$ (generating the $s_{d,l}$), and $H_d(x)$ (generating the $s_{p,m}$). The connection between the activation pattern and Eqs. (B1) and (B2) is illustrated in Fig. 7.

An active daughter of I, here labeled Y, has one fewer unexplored parent branch because of I, so

$$s_d = 1 + \sum_{l=1}^{\tilde{n}_p} s_{p,l} + \sum_{m=1}^{n_d} s_{d,m},$$

where $\tilde{n}_p$ ranges from 0 to $k-1$ and counts the parents other than I. Therefore, s_d is generated by H_p, which obeys the following self-consistent equation:

$$H_p(x) = xA_o(H_p(x))A_i(H_d(x)). \tag{B3}$$

Similarly, an active parent of I, here labeled X, has one fewer daughter branch to consider, so its cluster-size contribution is

$$s_p = 1 + \sum_{l=1}^{n_p} s_{p,l} + \sum_{m=1}^{\tilde{n}_d} s_{d,m},$$

where $\tilde{n}_d$ ranges from 0 to $k-1$, and counts the daughters other than I. Here, s_p is generated by H_d, which obeys the following self-consistent equation:

$$H_d(x) = xB_o(H_p(x))B_i(H_d(x)). \tag{B4}$$

The relationship between the three size-generating functions H_0, H_d, and H_p is illustrated in Fig. 8.

To summarize, H_p corresponds to the cluster size reached when arriving at a node from one of its parent

(a)   (b)   (c)

$$s = 1 + \sum_{l=1}^{n_d} s_{d,l} + \sum_{m=1}^{n_p} s_{p,m}$$

$$H_0(x) = x \prod_{l=1}^{n_p} H_d(x) \prod_{m=1}^{n_d} H_p(x)$$
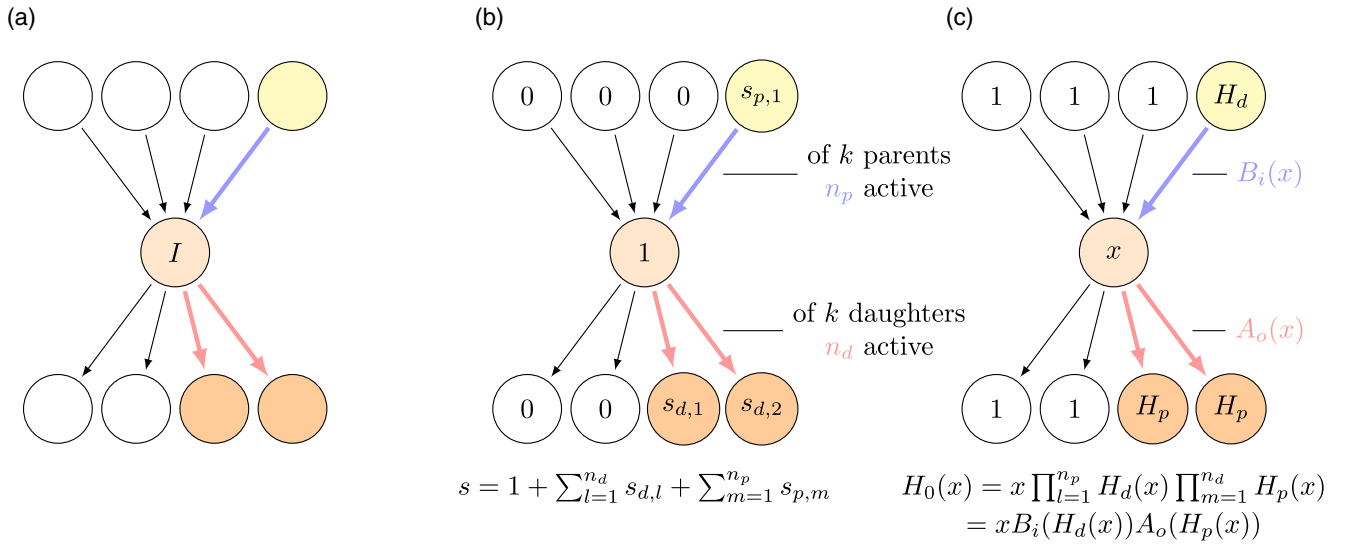$$= x B_i(H_d(x)) A_o(H_p(x))$$

FIG. 7.   Firing pattern example represented both as the sum of variables and the product of generating functions. (a) Example activation pattern beginning from a randomly selected initial active node I on a 4-regular network. Thick edges indicate connected active nodes. (b) Number of parent and daughter edges contributing to the cluster labeled by n_p and n_d, which are random variables that could vary from 0 to k. Nodes are labeled with their size contribution to the cluster—nodes that do not activate contribute zero, the initially considered node contributes one activation, while active parents and daughters contribute a random variable. (c) Same activity pattern labeled with the probability generating functions corresponding to each random variable.

branches, and H_d corresponds to the cluster size reached when arriving at a node from one of its daughter branches. The two pairs of generating functions (A_i, A_o) and (B_i, B_o) describe the number of active neighbors for H_p and H_d, respectively. In terms of the nodes labeled in Fig. 8, the neighbor generating functions and their corresponding probability mass functions are

$$A_i \leftrightarrow P(\tilde{n}_p \text{ active parents of } Y \text{ excluding}$$
$$I | Y \text{ active } \& I \text{ active}), \tag{B5}$$

$$A_o \leftrightarrow P(n_d \text{ active daughters of } Y | I \text{ active}), \tag{B6}$$

$$B_i \leftrightarrow P(n_p \text{ parents of } X | X \text{ active}), \tag{B7}$$
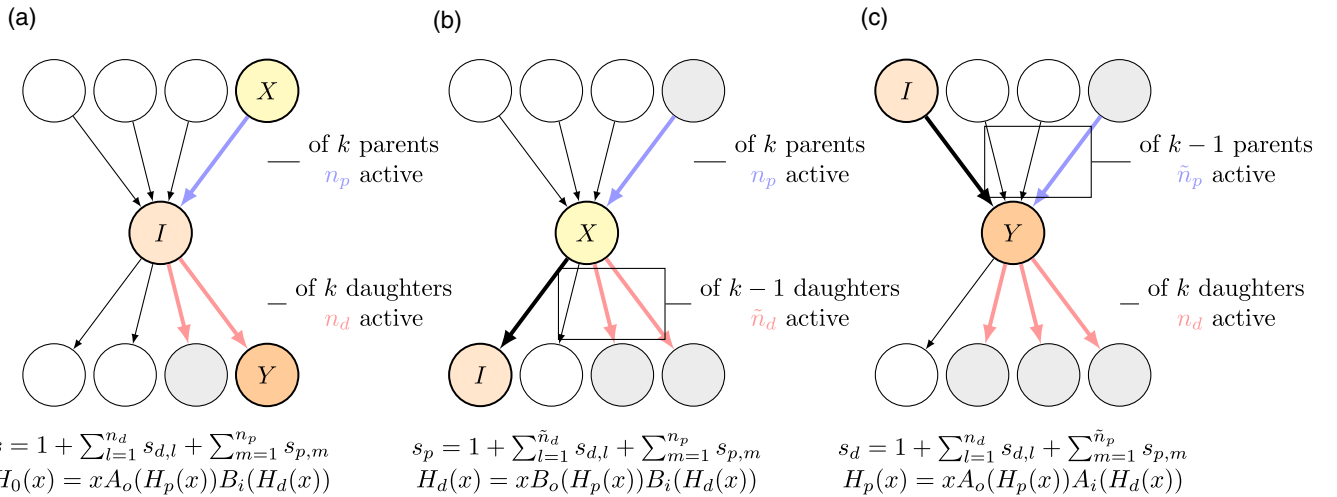
(a)   (b)   (c)

$$s = 1 + \sum_{l=1}^{n_d} s_{d,l} + \sum_{m=1}^{n_p} s_{p,m}$$
$$H_0(x) = x A_o(H_p(x)) B_i(H_d(x))$$

$$s_p = 1 + \sum_{l=1}^{\tilde{n}_d} s_{d,l} + \sum_{m=1}^{n_p} s_{p,m}$$
$$H_d(x) = x B_o(H_p(x)) B_i(H_d(x))$$

$$s_d = 1 + \sum_{l=1}^{n_d} s_{d,l} + \sum_{m=1}^{\tilde{n}_p} s_{p,m}$$
$$H_p(x) = x A_o(H_p(x)) A_i(H_d(x))$$

FIG. 8.   Example of the three size-generating functions. To illustrate the difference between H_0, H_d, and H_p, we consider the example firing pattern of Fig. 7 but centered on three different nodes, X, I, and Y. (a) Corresponding to H_0 and beginning from a randomly selected initial active node I on a 4-regular network. Thick edges indicate connected active nodes. Two neighboring active nodes, a parent and daughter (X and Y, respectively), are highlighted as corresponding to the other two generating functions. (b) Example activation pattern near X. Since one daughter connection leads to I, only $k-1$ are available for other connections. This restriction on daughters is why H_d differs from H_0. (c) Example activation pattern near Y. Since one of the parents of Y is I, only $k-1$ other parents need to be considered. Partly because of the restrictions on parents, H_p differs from H_0.

$B_o \leftrightarrow P(\tilde{n}_d$ active daughters of X excluding $I|X$ active). (B8)

As the giant component appears, the average cluster size diverges. Therefore, identifying the conditions under which the average cluster size diverges is a natural way to identify the critical line. For $q \leq q_c$, when $H_0(1) = H_p(1) = H_d(1) = 1$, the average cluster size is given by

$$\langle s \rangle_n = H'_0(1) = 1 + A'_o(1)H'_p(1) + B'_i(1)H'_d(1), \quad (B9)$$

where the subscript n denotes an average found by sampling randomly selected nodes instead of averaging over clusters. Now, since $A'_o(1)$ and $B'_i(1)$ correspond to the mean number of daughters and parents of the initial randomly selected node, quantities that are necessarily bounded above by the mean (in- or out-) degrees, $A'_o(1)$ and $B'_i(1)$, cannot diverge. Therefore, $\langle s \rangle_n$ can only diverge if $H'_p(1)$ or $H'_d(1)$ do. Using Eqs. (B3) and (B4), the following self-consistency relation for $H'_p(1)$ and $H'_d(1)$ (with $q \leq q_c$) can be obtained:

$$\begin{bmatrix} 1 - A'_o(1) & -A'_i(1) \\ -B'_o(1) & 1 - B'_i(1) \end{bmatrix} \begin{bmatrix} H'_p(1) \\ H'_d(1) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (B10)$$

Therefore, $H'_p(1)$ and $H'_d(1)$ diverge when the determinant of the above matrix is zero, i.e., when $0 = (1 - A'_o(1))(1 - B'_i(1)) - B'_o(1)A'_i(1)$. This condition will yield the critical line, when supplied with the PGFs for A and B.

1. Neighbor generating functions for k-regular networks

So far, we have been quite generic in developing the generating function H_0. To proceed further, we must supply $A_{i/o}$ and $B_{i/o}$ for a given network. For simplicity, we focus on the k-regular network. This approach allows us to develop expressions for Φ (the active fraction) and P_d (the probability that the daughter of an active site is activated in the next time step). The first quantity we need is the active fraction—the proportion of nodes activated in each time step. A randomly selected (not necessarily active) node will have m active parents with probability $\binom{k}{m}\Phi^m\bar{\Phi}^{k-m}$, as each parent is independent. With m parents, the probability of activation is $1 - \bar{p}\bar{q}^m$. Now, since the probability of activation for a random node is also Φ, we can write the self-consistent equation

$$\Phi = \sum_{m=0}^{k} \binom{k}{m}\Phi^m\bar{\Phi}^{k-m}(1 - \bar{p}\bar{q}^m)$$
$$= 1 - \bar{p}\overline{q\Phi}^k. \quad (B11)$$

It will be useful, when performing asymptotic analysis in the limit that $p \to 0$, to have a closed-form approximation for Φ. If we assume that $\Phi \ll 1$, we can truncate the expression $\bar{\Phi} = \bar{p}\{1 - kq\Phi + [k(k-1)/2]q^2\Phi^2 + \ldots\}$ to first or second order in Φ and solve for Φ, from which we obtain the first-order approximation

$$\Phi \approx \frac{p}{1 - kq} \quad (B12)$$

and the second-order approximation (choosing the positive root, since $\Phi > 0$)

$$\Phi \approx \frac{k\bar{p}q - 1 + \sqrt{1 - k^2\bar{p}^2q^2 - 2k\bar{p}q\overline{pq}}}{(k-1)k\bar{p}q^2}. \quad (B13)$$

For P_d, we have one active parent, and $k-1$ parents that are independently active with probability Φ. Hence,

$$P_d = \sum_{m=0}^{k-1} \binom{k-1}{m}\Phi^m\bar{\Phi}^{k-1-m}(1 - \bar{p}\bar{q}^{m+1})$$
$$= 1 - \bar{p}\,\bar{q}\,\overline{q\Phi}^{k-1},$$

and simplifying using Eq. (B11),

$$P_d = 1 - \frac{\bar{q}\,\bar{\Phi}}{\overline{q\Phi}}. \quad (B14)$$

Note that $\sigma = kP_d$ defines the branching ratio.

Now that we have both P_d and Φ, we can derive $A_{i/o}$ and $B_{i/o}$. The simplest ones to derive are $A_o(x)$ and $B_o(x)$ because they describe the number of activated daughters, and the activation of each daughter is independent of the others. We consider a single daughter, whose activation can be described by a single random variable $m \in \{0, 1\}$, with $m = 1$ only if the single daughter activates. The PGF corresponding to m is $C(x) = E(x^m) = \overline{P_d}x^0 + P_dx^1 = \overline{P_d} + P_dx$. If n is the number of activated daughters for a site with l available daughters, then $n = \sum_{i=1}^{l} m_i$ for m_l being independent and identically distributed (iid) Bernoulli variables generated by $C(x)$. Then, taking $l = k$ for A_o, we have $A_o(x) = E(x^n) = E(x^{\sum_{i=1}^{l} m_i}) = \prod_{i=1}^{k} E(x^{m_i}) = C(x)^k$, so

$$A_o(x) = (\overline{P_d} + P_dx)^k. \quad (B15)$$

For B_o, we have one fewer daughter from which to choose because we arrived at the node in question by means of one active daughter, so we take $l = k - 1$ to find

$$B_o(x) = (\overline{P_d} + P_dx)^{k-1}. \quad (B16)$$

Now, for A_i and B_i, we cannot treat the parents' activation as independent because it is conditional on the

knowledge that their daughter must activate and, in the case of A_i, also on the presence of other active parents.

Treating $B_i(x)$ first, we consider an active site (labeled X) that we arrive at by means of an active daughter (in Fig. 8, I). Therefore, we have no knowledge about the number of active parents, save for the fact that they successfully activated the node in question. Considering the probability mass function in Eq. (B7), Bayes' theorem allows us to write

$$P(n_p \text{ active parents of } X | X \text{ active})$$
$$= \frac{P(X \text{ active}|n_p \text{ active parents})P(n_p \text{ active parents})}{P(X \text{ active})}.$$

However, $P(X \text{ active}|n_p \text{ parents}) = 1 - \bar{p}\bar{q}^{n_p}$ by definition of the model [Eq. (1) of the main text], while the probability of n_p active parents, unconditioned on anything else, is just given by $P(n_p \text{ active parents}) = \binom{k}{n_p}\Phi^{n_p}\bar{\Phi}^{k-n_p}$. Lastly, the probability that X is active, conditioned on nothing else, is just the active fraction Φ. Thus,

$$P(n_p \text{ active parents of } X | X \text{ active})$$
$$= \frac{(1 - \bar{p}\bar{q}^{n_p})(\binom{k}{n_p}\Phi^{n_p}\bar{\Phi}^{k-n_p})}{\Phi}, \tag{B17}$$

which is exactly Eq. (A1). The generating function corresponding to B_i is therefore given by

$$B_i(x) = \sum_{n_p=0}^{k} P(n_p \text{ active parents of } X | X \text{ active})x^{n_p}$$

and can therefore be expressed as

$$B_i(x) = \frac{1}{\Phi}[(\bar{\Phi} + \Phi x)^k - \bar{p}(\bar{\Phi} + \Phi\bar{q}x)^k]. \tag{B18}$$

For $A_i(x)$, we consider a node Y that we arrived at from an active node (labeled I in Fig. 8), which is one of Y's parent branches. Here, $A_i(x)$ is the generating function for the number of additional active parents of Y. Considering the probability mass function in Eq. (B5) and applying Bayes' theorem, we find

$$P(\tilde{n}_p \text{ active parents of } Y \text{ excluding } I | Y \text{ active \& } I \text{ active})$$
$$= P(Y \text{ active}|I \text{ active \& } \tilde{n}_p \text{ other active parents of } Y)$$
$$\times \frac{P(\tilde{n}_p \text{ of the } k - 1 \text{ parents other than } I \text{ active})}{P(Y \text{ active}|I \text{ active})}.$$

Each of these probabilities is known:

$$P(Y \text{ active}|I \text{ active \& } \tilde{n}_p \text{ other parents of } Y \text{ active})$$
$$= 1 - \bar{p}\bar{q}^{\tilde{n}_p+1} \tag{B19}$$

by definition of the model [Eq. (1) of main text],

$$P(\tilde{n}_p \text{ of } k - 1 \text{ parents other than } I \text{ active})$$
$$= \binom{k-1}{\tilde{n}_p}\Phi^{\tilde{n}_p}\bar{\Phi}^{k-1-\tilde{n}_p}, \tag{B20}$$

and $P(Y \text{ active}|I \text{ active}) = P_d$. Hence,

$$P(\tilde{n}_p \text{ parents of } Y \text{ other than } I \text{ active}|Y \text{ active \& } I \text{ active})$$
$$= \frac{1}{P_d}(1 - \bar{p}\bar{q}^{\tilde{n}_p+1})(\binom{k-1}{\tilde{n}_p}\Phi^{\tilde{n}_p}\bar{\Phi}^{k-1-\tilde{n}_p}). \tag{B21}$$

Now, the generating function A_i is given by

$$A_i(x) = \sum_{\tilde{n}_p=0}^{k-1} x^{\tilde{n}_p} P(\tilde{n}_p \text{ active parents of } Y \text{ excluding}$$
$$I|Y \text{ active \& } I \text{ active}),$$

so after some algebra, we have

$$A_i(x) = \frac{1}{P_d}[(\bar{\Phi} + \Phi x)^{k-1} - \bar{p}\,\bar{q}\,(\bar{\Phi} + \Phi\bar{q}x)^{k-1}].$$

This concludes the calculation of the four generating functions $A_{i/o}$ and $B_{i/o}$ for the k-regular network. These calculations can also be conducted for other random networks, although the calculation is more technically involved when the in-degree can vary or correlations exist between the in- and out-degrees.

In summary, and in terms of Φ and P_d, the PGFs $A_{i/o}$ and $B_{i/o}$ for the k-regular network may be expressed as

$$A_o(x) = (\overline{P_d} + P_d x)^k, \tag{B22}$$

$$B_o(x) = (\overline{P_d} + P_d x)^{k-1}, \tag{B23}$$

$$A_i(x) = \frac{1}{P_d}[(\bar{\Phi} + \Phi x)^{k-1} - \bar{p}\,\bar{q}\,(\bar{\Phi} + \Phi\bar{q}x)^{k-1}], \tag{B24}$$

and

$$B_i(X) = \frac{1}{\Phi}[(\bar{\Phi} + \Phi x)^{k-1} - \bar{p}(\bar{\Phi} + \Phi\bar{q}x)^k]. \tag{B25}$$

2. Observables from the generating function

Here, we summarize how to extract observables, such as the size fraction of the giant component g, susceptibility χ, and cluster distribution $P_c(s)$ from the generating function $H_0(x)$. Practically speaking, we solve Eqs. (B3) and (B4) self-consistently for $H_d(x)$ and $H_p(x)$ via a Newton-Raphson scheme for a given set of model parameters p,

q, and x. With $H_d(x)$ and $H_p(x)$ in hand, we can insert these into Eq. (B2) and obtain $H_0(x)$.

The first quantity we can obtain from $H_0(x)$ is the fraction of nodes involved in finite clusters, which is just $H_0(1) = \sum_s p(s) = \bar{P}_\infty$, where $\bar{P}_\infty$ denotes the probability of belonging to the infinite cluster. So the giant component fraction g, the fraction of all nodes at all times that are part of the infinite cluster, is just $g = \Phi(1 - H_0(1))$. For the susceptibility $\chi = \langle s^2 \rangle = \sum s^2 P_c(s)$, we must make the distinction between the cluster-size distribution $P_c(s)$ [for numerical simulations, reported simply as $P(s)$] and the per-node cluster-size distribution $P_n(s)$. The latter describes the cluster sizes observed by sampling random active nodes and is directly calculated by the generating function approach or accessed by simulating avalanches on the infinite lattice. Clearly, $P_n(s) = As P(s)$, for a normalization factor A. Since $\sum P(s) = 1$, $A = \int_0^1 (1/x) H_0(x) dx$. So, $\chi = (1/A) \sum s P_n(s) s = (\langle s \rangle_n / A) = [H_0'(1)/\int_0^1 H_0(x) dx]$. Of course, we can directly access $P_n(s)$ by using $P_n(s) = (1/s!)[d^s H_0(x)/dx^s]|_{x=0}$. As was pointed out in Ref. [35], numerically evaluating this derivative for large s is most easily accomplished via a contour integral

$$\left.\frac{d^s H_0(x)}{dx^s}\right|_{x=0} = \frac{1}{2\pi i} \oint \frac{H_0(z) dz}{z^{s+1}}, \qquad (B26)$$

on the circle $z = e^{i\phi}$ for $\phi \in [0, 2\pi]$. Note that z^d becomes highly oscillatory at large d, so convergence of this integral can be improved via standard numerical techniques for oscillatory integrals [91]. The cluster probability distribution can then be accessed as $P(s) = (1/As)P_n(s)$.

3. Phase diagram for the k-regular network

We can study the divergence of $\chi \sim \langle s \rangle_n$ by solving Eq. (B10) and inserting the solution into Eq. (B9) to obtain

$$\langle s \rangle_n = \frac{1 - \frac{\sigma}{k}(\sigma - \sigma_m)}{(1 - \sigma)^2 - \frac{k-1}{k}\sigma\sigma_m}, \qquad (B27)$$

where $\sigma_m = A_i'(1) = (k-1)P_{p1} = (k-1)(\Phi/P_d)[1 - (\overline{P_d}^2/\bar{\Phi})]$ is the expected number of other active parents, to an active node with one already known parent. In other words, σ_m describes the rate of merging of initially independent clusters. Clearly, $\langle s \rangle_n$ diverges when $k(1 - \sigma)^2 - (k-1)\sigma\sigma_m = 0$ [Eq. (3) of the main text]. We could also arrive at this result by setting the determinant of Eq. (B10) to zero. A reparametrization that will be convenient when considering the correlation length is to replace σ_m with $\Upsilon = [(k-1)\sigma_m/k\sigma]$, meaning that the critical line diverges when
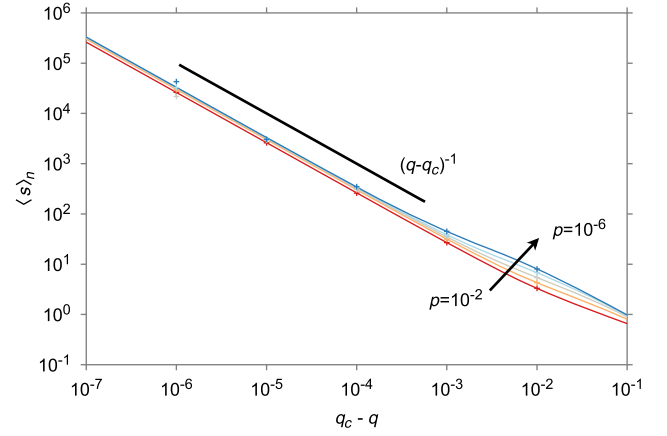
$$(1 - \sigma)^2 = \sigma^2 \Upsilon. \qquad (B28)$$

FIG. 9. Average cluster size for k-regular networks approaching criticality. Numerical simulations (points represent the mean of 10^6 realizations) on an infinite 10-regular graph yield good agreement with analytical predictions (lines) for $\langle s \rangle_n$.

The set of (p_c, q_c) that causes this divergence defines a critical line [see Fig. 5(a) in the main text].

Solving $0 = (1 - \sigma)^2 - \sigma^2 \Upsilon$ for q, and assuming $\Phi \ll 1$ (as occurs in the $p \ll 1$ limit with $q < q_c$) yields $(1/k) - p_1 = (k - 1/k^2)\sqrt{2k - 1}\sqrt{\Phi}$. Inserting the first-order closed-form approximation for $\Phi \ll 1$ [Eq. (B12)] into the solution for q yields the small p expansion for the phase curve [Eq. (4) in the main text]

$$\left(\frac{1}{k} - q\right)^3 = \frac{(k-1)^2(2k-1)}{k^5} p. \qquad (B29)$$

The average cluster size $\langle s \rangle_n$ (and therefore susceptibility χ) diverges for (p, q) near points on the critical line (p_c, q_c) as $\langle s \rangle_n \sim |p_c - p|^{-\gamma}$ (for $q = q_c$) and $\langle s \rangle_n \sim |q_c - q|^{-\gamma}$ (for $p = p_c$) with $\gamma = 1$, which is a direct consequence of the fact that the numerator and denominator of Eq. (B27) cannot both be simultaneously zero (except for the degenerate $q = 1$ case). Hence, the behavior near the critical line will depend only on how the denominator $f(p, q) = (1 - \sigma)^2 - (k - 1/k)\sigma\sigma_m$ scales near its zero p_c, q_c. As $(\partial f/\partial p) \neq 0$ and $(\partial f/\partial q) \neq 0$ at (p_c, q_c), the Taylor series approximation $f(p, q) \approx (\partial f/\partial p)(p - p_c) + (\partial f/\partial q)(q - q_c)$. Choosing $p = p_c$ or $q = q_c$ immediately yields the power-law scaling exponent $\gamma = 1$. This divergence can be visualized in Fig. 9.

4. Giant component

The giant component fraction g is given by $g = \Phi[1 - H_0(1)] = \Phi(1 - H_d(1)[\overline{P_d} + P_d H_p(1)])$. At the critical point, $H_p(1) = H_d(1) = 1$. So for $\delta = q - q_c \ll 1$, we have that $g \approx [\Phi P_d(\partial H_p/\partial q) + \Phi(\partial H_d/\partial q)]\delta$. Since both $H_d(1)$ and $H_p(1)$ are strictly decreasing functions of q, $g \sim (q - q_c)$, identifying the critical exponent $\beta = 1$.

TABLE III. Critical exponents for the k-regular network, beyond those reported in Table I of the main text, where $\delta q = |q_c - q|$ and $\delta q_{DP} = |q_{c,DP} - q|$, where $q_{c,DP}$ denotes the directed-percolation critical point (i.e., at $p = 0$). Exponents related to temporal dynamics (i.e., α and $1/\sigma\nu z$) reflect numerical observations from Fig. 17, while all other exponents are analytically determined.

Exponent	Quantity	This work	Literature
γ_{DP}	$\chi = \langle s^2 \rangle_c \sim \delta q^{-\gamma_{DP}}$ for $(\delta q/p) \gg 1$	3	3 [37] (using $\gamma = (3 - \tau/\sigma)$)
γ	$\chi \sim \delta q^{-\gamma}$ for $(\delta q/p) \ll 1$	1	1 [38]
α_{DP}	$P(T) \sim T^{-\alpha_{DP}}$ for $T < T_m$	≈ 2	2 [37] (using $\alpha = \delta + 1$)
α	$P(T) \sim T^{-\alpha}$ for $T > T_m$	≈ 7	$\cdots$
$1/\sigma\nu z_{DP}$	$T \sim s^{1/\sigma\nu z_{DP}}$ for $s < s_m$	≈ 2	2 [37]
$1/\sigma\nu z$	$T \sim s^{1/\sigma\nu z}$ for $s > s_m$	≈ 4	-
$\nu_{\parallel}$	$\xi_{\parallel} \sim \delta q^{-\nu_{\parallel}}$	1/2	1/2 (isotropic percolation) [38]
$1/\sigma_{DP}$	$s_m \sim \delta q_{DP}^{-1/\sigma_{DP}}$	2	2 [37]
$1/\sigma$	$P(s) \sim s^{-\tau} G(s/s_\xi)$ for $s > s_m$ and $G(x \gg 1) \to 0$ then $s_\xi \sim \delta q^{-(1/\sigma)}$	2	2 [38]

APPENDIX C: MERGELESS AVALANCHES

The mergeless clusters are exactly those clusters with one root. The number of configurations of singly rooted clusters of size s is given by the Fuss-Catalan numbers $C_s^{(k)} = \frac{1}{(k-1)s+1}\binom{ks}{s}$, which count the number of incomplete k-ary trees with s vertices [92]. Such a tree has perimeter (unoccupied branches) of length $t = (k-1)s + 1$. Nodes are included in the tree with probability

$$P_{d1} = \bar{\Phi}^{k-1}(1 - \bar{p}\,\bar{q}), \qquad (C1)$$

denoting the probability that a given daughter node is activated while having exactly one parent. The excluded nodes on the perimeter occur with probability $\bar{P}_d$, which is the probability of not activating, despite having an active parent. Hence, the probability of observing a mergeless cluster of size s is given by $P(s) = C_s^{(k)} P_{d1}^{s-1} \bar{P}_d^{\,t}$. Applying Stirling's approximation, we find

$$P(s) \sim s^{-3/2}\left(\frac{P_{d1}k}{((1-\frac{1}{k})/\bar{P}_d)^{k-1}}\right)^s = s^{-3/2}e^{-s/s_m}, \quad (C2)$$

where the characteristic mergeless size is

$$s_m = -1/\log\left[kP_{d1}\left(\frac{\bar{P}_d}{1 - \frac{1}{k}}\right)^{k-1}\right]. \qquad (C3)$$

In the limit of $p \ll 1$ and $(1/k) - q \ll 1$, expressing P_d [Eq. (B14)] and P_{d1} [Eq. (C1)] in p, q, and Φ, and again using the closed-form approximation for Φ [Eq. (B12)], we find to lowest order in p and $[(1/k) - q]$,

$$s_m^{-1} \approx \frac{(2k-1)(k-1)}{k^2}p\left(\frac{1}{k} - q\right)^{-1} + \frac{k^3}{2(k-1)}\left(\frac{1}{k} - q\right)^2, \qquad (C4)$$

and if we apply Eq. (4) to observe how the cutoff scales on the critical line, we find

$$s_m \approx \frac{2(k-1)}{3k^3}\left(\frac{1}{k} - q\right)^{-2}, \qquad (C5)$$

which is Eq. (5) from the main text. Since $q_{c,DP} = (1/k)$ and the relation $s_m \sim (q_{c,DP} - q)^{-1/\sigma^{DP}}$ defines the directed-percolation exponent σ^{DP}, we have also recovered the usual directed-percolation exponent $\sigma^{DP} = \frac{1}{2}$ (cf. Table III).

APPENDIX D: WIDOM LINE

In equilibrium critical points, divergence in the correlation length is associated with a divergence in the susceptibility of the order parameter to an infinitesimal application of an external field. In directed percolation, the order parameter is Φ. The susceptibility measures the activity
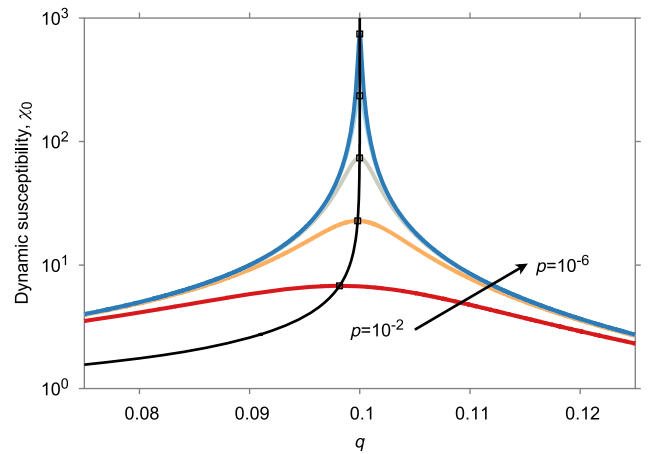

FIG. 10. Dynamic susceptibility for various p and q as calculated for an infinite 10-regular graph. For each p, there is a corresponding q that maximizes the susceptibility. These maxima are labeled by the squares, and they fall on the Widom line.

of the system in response to an external stimuli. We can imagine that the external stimuli is an infinitesimal increase in the average spontaneous activity of the system, and hence we can define the dynamic susceptibility as $\chi_0 \equiv (\partial \Phi / \partial p)$. So, using Eq. (B11), we find

$$\chi_0 = \overline{\Phi q}^k + qk\bar{p}\overline{\Phi q}^{k-1}\chi_0,$$

$$\Rightarrow \chi_0 = \frac{\bar{p}\overline{\Phi q}^{k+1}}{\bar{p}(\overline{\Phi q} - qk\bar{p}\overline{\Phi q}^k)},$$

and simplifying with Eq. (B11), we obtain

$$\chi_0 = \frac{\bar{\Phi}\,\overline{\Phi q}}{\bar{p}(\overline{\Phi q} - kq\overline{\Phi})}. \tag{D1}$$

In the limit $p \to 0$ with $\Phi \to 0$ and $q \to 1/k$, χ_0 is, asymptotically, $\chi_0 \sim (1/k)[(1/k) - q]^{-1}$ and therefore diverges at the directed-percolation critical point $q = (1/k)$ and $p = 0$. This susceptibility has been studied in the context of neural systems, where the mixing of initiation and spreading timescales means χ_0 no longer diverges (cf. Fig. 10) but instead is maximized on a quasicritical "Widom" line, where the fluctuations var$[\Phi(t)]$ are also maximized [42].

APPENDIX E: PHASE-CURVE SCALING

In directed percolation, there are several indicators of the critical point. The mean cluster size diverges, the branching ratio is 1, and the dynamic susceptibility diverges. However, with the introduction of spontaneous activation, it is clear that these indicators no longer agree [cf. Fig. 5(a) of the main text, or Fig. 11]. In fact, the branching ratio is no longer a clear signal because independent streams of activity can merge together and nodes can spontaneously
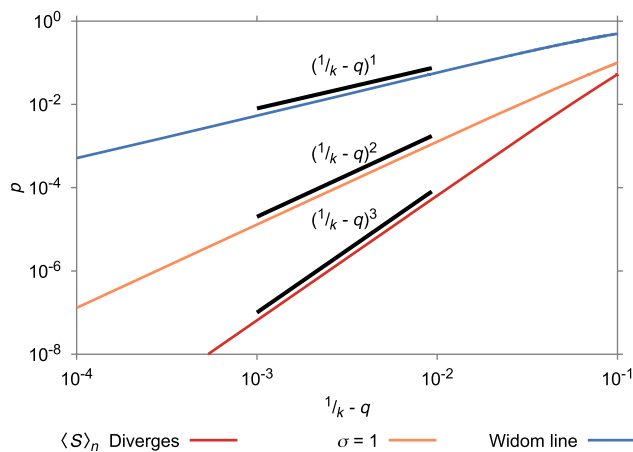

FIG. 11. Power-law scaling of the critical and quasicritical lines near the directed-percolation limit. From bottom to top, we show the critical line, $\sigma = 1$, and the Widom line in the $p \to 0$ limit.

activate. Meanwhile, the dynamic susceptibility no longer diverges but instead attains a maximum on a set of (p, q) that define the Widom line.

Although the Widom line, the unity branching ratio ($\sigma = 1$) line, and the line of diverging cluster size all agree as $p \to 0$, they obey different power laws in their approach to that point (Fig. 11). In this section, we derive the different scalings associated with these critical and quasi-critical lines.

The scaling for the $\sigma = 1$ line is given by $[(1/k) - q]^2 \sim p$, which can be seen by solving Eq. (B14) for q and using the closed form for Φ [Eq. (B12)], which immediately yields $(k^2/k - 2)[(1/k) - q]^2 \approx p$ on the $\sigma = 1$ line.

As for the Widom line, by setting $(\partial \chi_0 / \partial q) = 0$ and applying some simple algebraic manipulation, the Widom line can be found to consist of the q and p satisfying $0 = 1 - kq\bar{\Phi}^2 - 2\Phi + q\Phi^2$. The first-order approximation for Φ given by Eq. (B12) is poor in the vicinity of the Widom line (after all, it is in the vicinity of the point of maximum susceptibility in Φ), so a second-order approximation for Φ [Eq. (B13)] is necessary. With this approximation, the Widom line becomes (upon expansion around $p = 0$ and $q = (1/k)$) $p \approx (k/2)[(1/k) - q]$.

The critical line was previously shown [Eq. (4)] to obey the scaling $[(1/k) - q]^3 \sim p$. These scalings are illustrated in Fig. 11.

APPENDIX F: PHASE DIAGRAM AND SCALING COLLAPSE ON SMALL-WORLD NETWORKS

By numerically determining the critical point for different p values, we can build a critical line for the small-world networks. In the main text, we showed that the avalanche distribution $P(s)$ exhibits a scaling collapse when assuming $s_m \sim p^{-2/3}$, for a small-world network with a rewire probability of 10^{-2}. This result was somewhat surprising, as the small-world network still exhibits a large number of recurrent connections and is very nearly a circulant graph. However, for a lower rewire probability of 10^{-3}, recurrent connections play an even larger role, and $p^{-2/3}$ does not provide such a robust scaling. Instead, we find the collapse is best for $p^{-0.75}$ (cf. Fig. 12), which can be understood by considering the root-size distribution at the critical point, as is done in Fig. 12(c). Clearly, the characteristic merging size s_m scales as $s_m \sim p^{-0.75}$.

Interestingly, this method also allows us to estimate the directed-percolation $1/\sigma_{DP}$ exponent for the small-world network. Given a phase line that scales as $p \sim (q_{c,DP} - q)^a = \delta_{DP}^a$ and a curve collapse effected by p^b, and given the fact that we expect the curve collapse to scale as $s_m \sim \delta_{DP}^{1/\sigma_{DP}}$, we have the scaling relation $1/\sigma_{DP} = ab$. The phase diagram for this small-world network [Fig. 12(b)] relates $p \sim (q_{c,DP} - q)^{3.0908}$ for $q_{c,DP} \approx 0.11533$, which implies $\sigma_{DP} \approx \frac{1}{3.09 \times 0.75} = 0.43$ for the small-world network
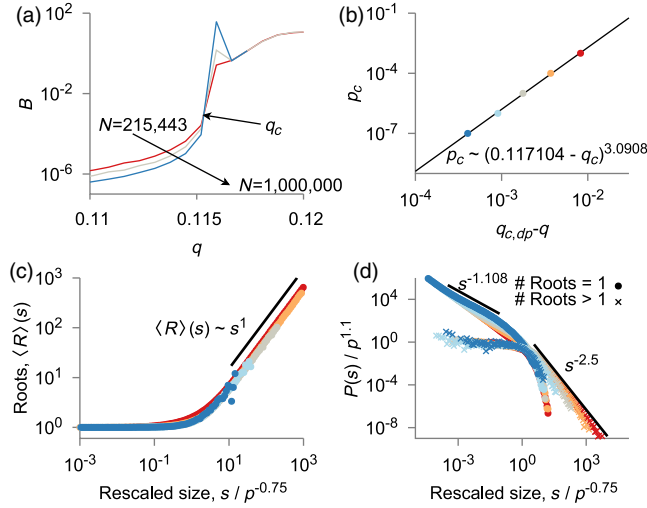
FIG. 12. Critical line and avalanche scaling of small-world network with a low rewire probability. The phase diagram and curve collapse along that phase line for $N = 10^6$ small-world networks with rewire probability 10^{-3}. (a) Ratio $B = g(\langle s^2 \rangle_c / \langle s \rangle_c^2)$ for $p = 10^{-5}$ identifying the critical point $q_c = 0.115333 \pm 0.000010$. (b) Numerically derived critical line represented with symbols. The black line is the nonlinear least-squares fit to the data. (c) Average number of roots, which exhibits a transition that scales with $s_m \sim p^{-0.75}$. (d) Avalanche distribution partitioned into merging and mergeless avalanches is collapsed when rescaled by s_m.

with a rewire probability 10^{-3}. When the shortcut density is low, the small-world network is approximately a one-dimensional circulant graph, which suggests we should use the $(1 + 1)$-dimensional directed-percolation exponents. Our result of $\sigma_{DP} \approx 0.43$ compares reasonably well with the $\sigma_{DP} = 0.391$ reported in the literature [37]. Similarly, we should identify the avalanche exponent $\tau_{DP} = 1.108$, which matches well with our numerical results (cf. Fig. 12).

APPENDIX G: CORRELATION LENGTH

In the main text, we introduce the pair-connectedness function $c(d, t)$. Typically in directed percolation, the pair-connectedness function can be used to define the correlation lengths $\xi_\parallel$ and $\xi_\perp$ by means of $c_{\parallel/\perp}(d) \sim \exp[-d/\xi_{\parallel/\perp}]$ (where $c_{\parallel/\perp}$ measures connectedness in the temporal or spatial directions, respectively). While this approach works well for finite-dimensional lattices, the loopless k-regular graphs we consider are effectively *infinite* dimensional, as the perimeter grows exponentially with network distance. As a consequence, the diverging (spatial) characteristic network length $\ell_\perp$ defined by $c(2d, 0)$ [in Eq. (G1) as in the main text] can differ from the correlation length ξ defined by the radius of gyration [38]. In particular, on the Bethe lattice, $\ell \sim |\delta q|^{-1}$, while $\xi \sim |\delta q|^{-1/2}$ [38].

We will show analytically that the perpendicular correlation length $\ell_\perp$ corresponding to the decay of $c(d, 0)$ diverges on the critical line.

1. Divergence of perpendicular characteristic length

We can derive the divergence of the perpendicular characteristic length $\ell_\perp$ by computing $c(2d, 0)$ and identifying the scaling

$$c(2d, 0) \sim \exp[-d/\ell_\perp]. \tag{G1}$$

We consider $2d$ because when a daughter branch is followed, t advances by 1, while a parent branch decreases t by 1. However, $\Delta t = 0$, so the number of parent and daughter branches must both be equal and their sum even.

Thus, not all routes with d daughters and d parent connections are equally likely. For instance, whenever a parent connection follows a daughter connection, the route requires that two initially independent avalanches merge at that point. It turns out that the most convenient way to compute $c(2d, 0)$ is to sum over collections of routes that have a fixed number of merges. The number of routes with m merges is given by

$$\left[\binom{d}{m} + \frac{1}{k-1} \binom{d-1}{m-1} \right]^2 k^{2d-1} \left(\frac{k-1}{k} \right)^{2m}. \tag{G2}$$

Given a sequence of parent-daughter network hops, the combinatorial factor $\binom{d}{m}$ counts the number of ways that the parent-daughter network hops could be rearranged without altering the number of merges. A network hop that follows a network hop of the same kind (i.e., a parent hop following a parent hop, or a daughter hop following a daughter) contributes k possible paths. Every time a change in direction occurs (i.e., a daughter followed by a parent or a parent by a daughter), only $k - 1$ links are available because we exclude the link by which we arrived at the node. The factor of $k^{2d-1}(k - 1/k)^{2m}$ captures the number of possible paths, given a sequence of daughter-parent network hops. A correction of $1/k - 1\binom{d-1}{m-1}$ is required to account for those paths with one or two fewer changes in direction (a boundary condition effect). A daughter or parental connection occurs with weight P_d or $P_p = P_d$, except when a parental connection follows a daughter connection, when it instead contributes P_{p1}. So, since the number of merges could range from 0 to d, we can write

$$c(2d, 0) = \sum_{m=0}^{d} \left(\left[(k-1) \binom{d}{m} + \binom{d-1}{m-1} \right]^2 \right.$$
$$\left. \times \left(\frac{k-1}{k} \right)^{2m-1} \frac{(kP_d)^{2d}}{k^2} \left(\frac{P_{p1}}{P_d} \right)^m \right).$$

This expression is compared to simulations on the infinite lattice in Fig. 13 and has the closed-form expansion
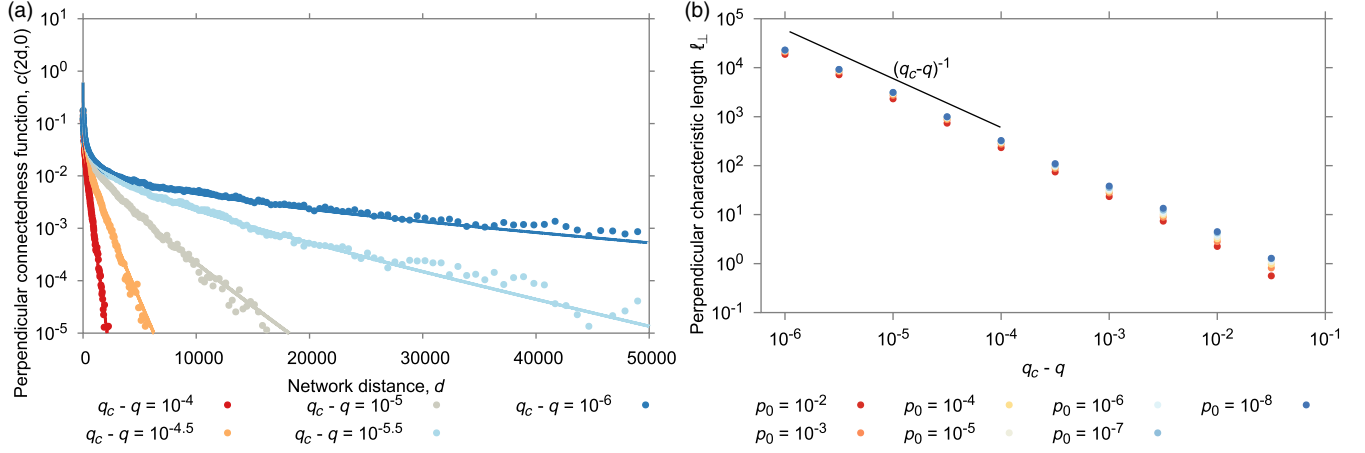
FIG. 13. Perpendicular characteristic length near criticality on infinite k-regular networks. (a) Simultaneous (perpendicular) path connectedness function with an exponentially decaying tail, with the exact form predicted analytically. Solid lines are analytical results, while symbols are numerical simulations on infinite 10-regular networks (averaged over 20,000,000 clusters), simulated at $p_0 = 10^{-3}$. (b) Perpendicular characteristic length diverging with a power law as $\ell_\perp \sim (q_c - q)^{-1}$.

$$c(2d, 0)\frac{k\sigma - (k-1)^2 P_{p1}}{\sigma^d}$$
$$= -(2(\sigma - (k-1)P_{p1})\,_2F_1(1-d, -d; 1; \Upsilon)$$
$$+ ((k+1)\sigma + kP_{p1})\,_2F_1(1-d, 1-d; 1; \Upsilon)),$$

where $_2F_1$ denotes the Gauss hypergeometric function, and $\sigma = kP_d$ and $\Upsilon = [(k-1)^2 P_{p1}/k^2 P_d] = [(k-1)\sigma_m/k\sigma]$ as before. Since in the limit of large d, $[_2F_1(1-d, -d; 1; \Upsilon)/\,_2F_1(1-(d-1), -(d-1); 1; \Upsilon)] = [_2F_1(1-d, 1-d; 1; \Upsilon)/\,_2F_1(1-(d-1), 1-(d-1); 1; \Upsilon)]$, we need only treat one of them asymptotically. By way of Kummer's 24 solutions [93], we have that $_2F_1(1-d, -d; 1; \Upsilon) = (1-\Upsilon)^d\,_2F_1(d, -d; 1; \Upsilon/(1-\Upsilon))$. Using an identity from Wilson [94], we have $_2F_1(1-d, -d; 1; \Upsilon) \sim [(1+\sqrt{\Upsilon})^d/\sqrt{d}]$, so in the limit of large d,

$$\frac{c(2(d+1), 0)}{c(2d, 0)} \approx (\sigma(1 + \sqrt{\Upsilon}))^2, \qquad (G3)$$

which yields $\ell_\perp = [-1/2\log(\sigma(1 + \sqrt{\Upsilon}))]$. Now, as $(\partial/\partial q)\sigma(1 + \sqrt{\Upsilon}) > 0$ at $q = q_c$ [which can be seen by using Eqs. (4) and (B12)], we can write $\sigma(1 + \sqrt{\Upsilon}) \approx 1 + c(q - q_c) + \mathcal{O}((q - q_c)^2)$ by way of a Taylor expansion, implying $\ell_\perp \sim (q_c - q)^{-1}$. Finally, it is clear that the divergence in correlation length occurs precisely when $\sigma^2(1 + \sqrt{\Upsilon})^2 = 1$. Simple algebra shows that this is equivalent to $\sigma^2\Upsilon = (1 - \sigma^2)$, which is the same critical line derived by considering the divergence of the average cluster size [see Eq. (B28)].
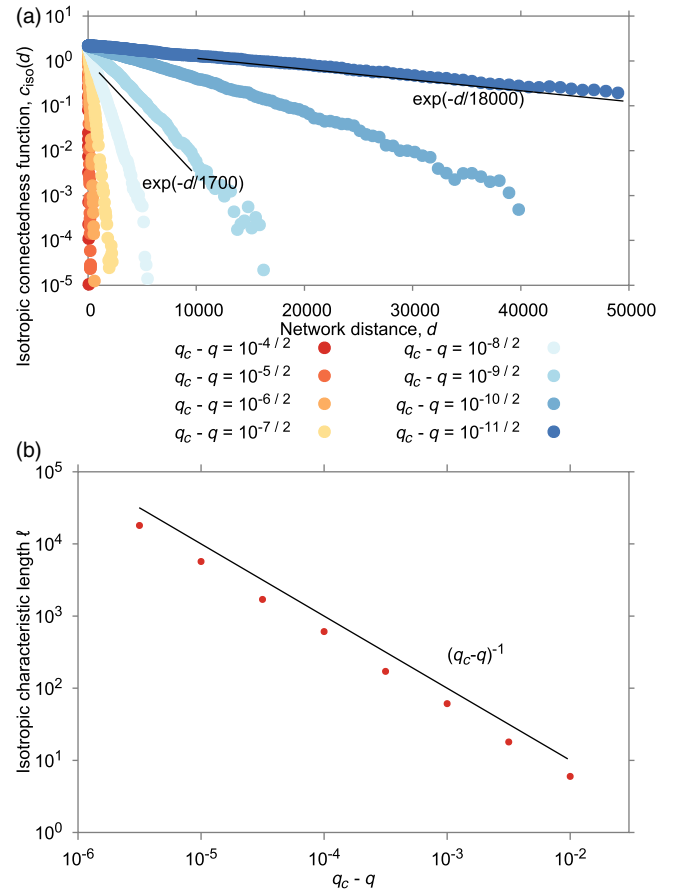

FIG. 14. Isotropic characteristic network length near criticality on infinite k-regular networks. (a) Isotropic path connectedness function $c_{iso}(d)$ for all paths of length d decaying exponentially, which is the result of simulations of infinite 10-regular networks, simulated at $p = 10^{-3}$. (b) Corresponding isotropic characteristic network length ℓ, which diverges with a power law of $(q_c - q)^{-1}$.
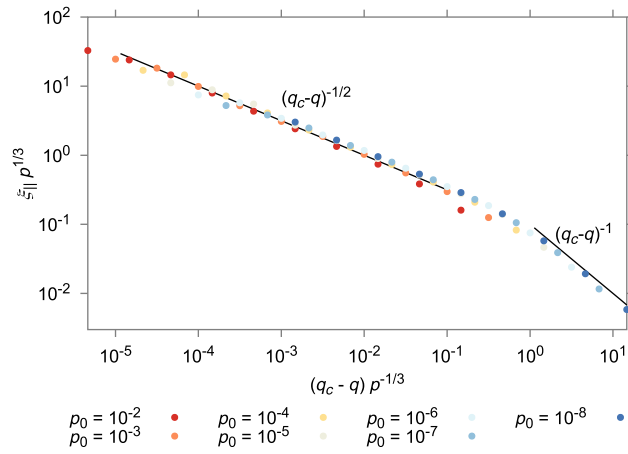
FIG. 15. Timelike correlation length $\xi_\parallel$ on infinite k-regular networks, collapsed for different values of spontaneous activity p, exhibiting a divergence with $\xi_\parallel \sim (q_c - q)^{-1}$.
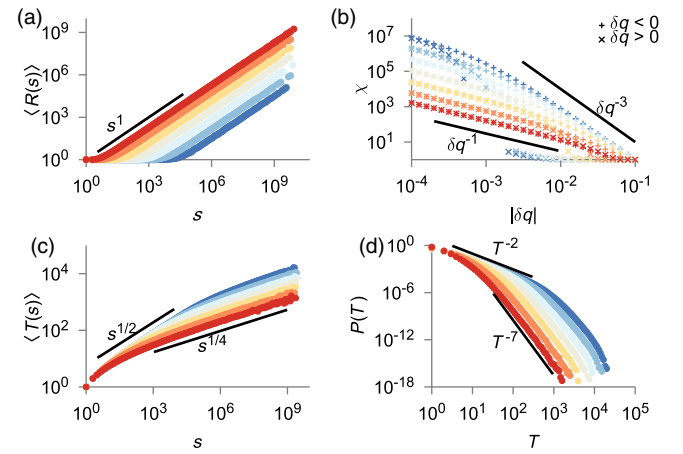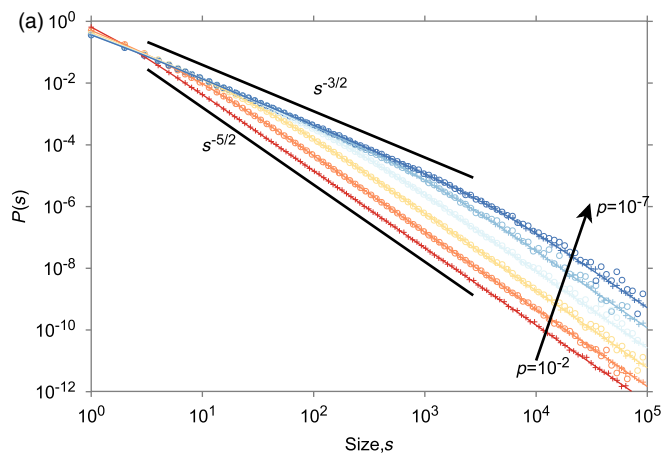

FIG. 17. Exponent transitions for critical avalanches, without the curve collapse presented in the main text. Exponent transitions for 10-regular networks, without the rescaling presented in Fig. 5 of the main text. Data are for $p = 10^{-2}$ to $p = 10^{-8}$. (a) Average number of roots for avalanches of a given size, as simulated on infinite networks. (b) Analytically determined susceptibility χ. (c) Avalanche duration-size relation. (d) Avalanche duration distribution.

2. Isotropic characteristic length

We have considered a characteristic spatial length $\ell_\perp$, in analogy to the directed-percolation perpendicular correlation length. Note that $\ell_\perp$ is anisotropic, which may appear to make it unrelated to the isotropic correlation length of undirected percolation. However, the presence of a diverging anisotropic characteristic length implies that any isotropic characteristic length will also diverge. To illustrate this, consider the isotropic characteristic length ℓ defined by the mean number of sites active after d (isotropic) network hops away from an active node, i.e., $c_{\mathrm{iso}}(d) = \sum_{t=-d}^{d} c(d,t) \sim \exp[-d/\ell]$. Since this sum contains $c(d,0) \sim \exp[-d/(2\ell_\perp)]$, which tends to a constant as $q \to q_c$, we know that ℓ also diverges in the same limit, as can be seen in Fig. 14.

3. Temporal correlation lengths

Directed-percolation phase transitions have a diverging temporal length scale. In pure directed percolation, we can directly measure this by considering the decay in activity from an active site, $c(t,t) = \sigma^t \sim \exp[-t/\xi_t]$, which counts the number of direct descendants from an active site. Then, $\xi_t = -1/\log[\sigma]$, implying that this correlation length diverges on the $\sigma = 1$ line, with $\xi_t \sim |\delta q|^{-1}$, consistent with $\nu_\parallel$ as expected in mean-field directed percolation.

However, $c(t,t)$ can only have t network hops, which restricts it to including only direct *causal* descendants. Nonetheless, activity at a node can still be *correlated* with
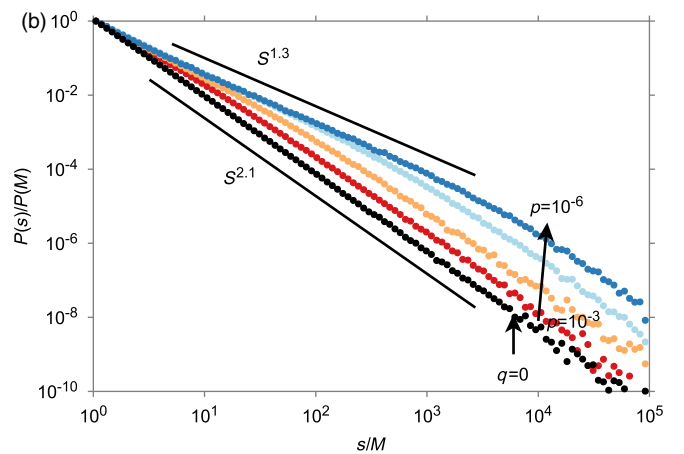

FIG. 16. Unscaled critical avalanche distributions. (a) Avalanche distributions for 10-regular networks as in Fig. 3(e). Solid lines are analytical $P(s)$ determined from the generating function, while the crosses and circles are simulations on infinite and finite ($N = 10^7$) networks. (b) Avalanche distributions for HMN networks as in Fig. 3(f) of the main text.

activity at much later times, even if its direct descendants are extinct, because of activity from its progenitors. We define the correlation length $\xi_\parallel$ by $c(t) \sim \sum_d c(d,t) \sim \exp[-t/\xi_\parallel]$. We measure the evolution of $\xi_\parallel$ in infinite k-regular network simulations and find two power-law regimes (cf. Fig. 15), separated by $|\delta q| \approx p^{1/3}$. For low values of p, we have $\nu_\parallel = 1$, as in directed percolation. However, for sufficiently small δq, we have $\nu_\parallel = 1/2$, as in isotropic percolation.

APPENDIX H: ADDITIONAL FIGURES

In this section, we include several figures to supplement the main text. Figure 16 contains unscaled avalanche distributions for the hierarchical modular networks and 10-regular networks corresponding to Fig. 3. Similarly, Fig. 17 consists of the exponent transitions for 10-regular graphs without the exponent transitions presented in Fig. 5.

[1] M. E. Newman, *Spread of Epidemic Disease on Networks*, Phys. Rev. E **66**, 016128 (2002).

[2] E. Kenah and J. M. Robins, *Second Look at the Spread of Epidemics on Networks*, Phys. Rev. E **76**, 036113 (2007).

[3] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Epidemic Processes in Complex Networks*, Rev. Mod. Phys. **87**, 925 (2015).

[4] J. C. Wierman and D. J. Marchette, *Modeling Computer Virus Prevalence with a Susceptible-Infected-Susceptible Model with Reintroduction*, Computational Statistics and Data Analysis **45**, 3 (2004).

[5] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili, *Theory of Rumour Spreading in Complex Social Networks*, Physica (Amsterdam) **374A**, 457 (2007).

[6] P. Crucitti, V. Latora, and M. Marchiori, *A Topological Analysis of the Italian Electric Power Grid*, Physica (Amsterdam) **338A**, 92 (2004).

[7] X. Wang, Y. Koç, R. E. Kooij, and P. Van Mieghem, *A Network Approach for Power Grid Robustness Against Cascading Failures*, in *2015 7th International Workshop on Reliable Networks Design and Modeling (RNDM)* (IEEE, New York, 2015), pp. 208–214.

[8] D. R. Chialvo, *Emergent Complex Neural Dynamics*, Nat. Phys. **6**, 744 (2010).

[9] N. Friedman, S. Ito, B. A. Brinkman, M. Shimono, R. L. 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).

[10] J. G. Orlandi, J. Soriano, E. Alvarez-Lacalle, S. Teller, and J. Casademunt, *Noise Focusing and the Emergence of Coherent Activity in Neuronal Cultures*, Nat. Phys. **9**, 582 (2013).

[11] E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. Chialvo, *Criticality in Large-Scale Brain FMRI Dynamics Unveiled by a Novel Point Process Analysis*, Front. Physiol. **3**, 15 (2012).

[12] S. Di Santo, P. Villegas, R. Burioni, and M. A. Muñoz, *Landau–Ginzburg Theory of Cortex Dynamics: Scale-Free Avalanches Emerge at the Edge of Synchronization*, Proc. Natl. Acad. Sci. U.S.A. **115**, E1356 (2018).

[13] O. Kinouchi and M. Copelli, *Optimal Dynamical Range of Excitable Networks at Criticality*, Nat. Phys. **2**, 348 (2006).

[14] M. Yaghoubi, T. de Graaf, J. G. Orlandi, F. Girotto, M. A. Colicos, and J. Davidsen, *Neuronal Avalanche Dynamics Indicates Different Universality Classes in Neuronal Cultures*, Sci. Rep. **8**, 3417 (2018).

[15] D. Dahmen, S. Grün, M. Diesmann, and M. Helias, *Second Type of Criticality in the Brain Uncovers Rich Multiple-Neuron Dynamics*, Proc. Natl. Acad. Sci. U.S.A. **116**, 13051 (2019).

[16] J. M. Beggs and D. Plenz, *Neuronal Avalanches in Neocortical Circuits*, J. Neurosci. **23**, 11167 (2003).

[17] A. Ponce-Alvarez, A. Jouary, M. Privat, G. Deco, and G. Sumbre, *Whole-Brain Neuronal Activity Displays Crackling Noise Dynamics*, Neuron **100**, 1446 (2018).

[18] B. Del Papa, V. Priesemann, and J. Triesch, *Criticality Meets Learning: Criticality Signatures in a Self-Organizing Recurrent Neural Network*, PLOS One **12**, e0178683 (2017).

[19] J. M. Beggs and N. Timme, *Being Critical of Criticality in the Brain*, Front. Physiol. **3**, 163 (2012).

[20] L. Cocchi, L. L. Gollo, A. Zalesky, and M. Breakspear, *Criticality in the Brain: A Synthesis of Neurobiology, Models and Cognition*, Prog. Neurobiol. **158**, 132 (2017).

[21] J. M. Beggs and D. Plenz, *Neuronal Avalanches Are Diverse and Precise Activity Patterns That Are Stable for Many Hours in Cortical Slice Cultures*, J. Neurosci. **24**, 5216 (2004).

[22] P. Moretti and M. A. Muñoz, *Griffiths Phases and the Stretching of Criticality in Brain Networks*, Nat. Commun. **4**, 2521 (2013).

[23] 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).

[24] L. Dalla 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).

[25] S.-S. Poil, R. Hardstone, H. D. Mansvelder, and K. Linkenkaer-Hansen, *Critical-State Dynamics of Avalanches and Oscillations Jointly Emerge from Balanced Excitation/Inhibition in Neuronal Networks*, J. Neurosci. **32**, 9817 (2012).

[26] J. P. Neto, F. P. Spitzner, and V. Priesemann, *A Unified Picture of Neuronal Avalanches Arises from the Understanding of Sampling Effects*, arXiv:1910.09984.

[27] A. Das and A. Levina, *Critical Neuronal Models with Relaxed Timescale Separation*, Phys. Rev. X **9**, 021062 (2019).

[28] R. V. Williams-Garcia, J. M. Beggs, and G. Ortiz, *Unveiling Causal Activity of Complex Networks*, Europhys. Lett. **119**, 18003 (2017).

[29] P. Van Mieghem and E. Cator, *Epidemics in Networks with Nodal Self-Infection and the Epidemic Threshold*, Phys. Rev. E **86**, 016116 (2012).

[30] H. Hinrichsen, *Non-Equilibrium Critical Phenomena and Phase Transitions into Absorbing States*, Adv. Phys. **49,** 815 (2000).

[31] M. Yanuka, *The Mixed Bond-Site Percolation Problem and Its Application to Capillary Phenomena in Porous Media*, J. Colloid Interface Sci. **134,** 198 (1990).

[32] H. F. Song and X.-J. Wang, *Simple, Distance-Dependent Formulation of the Watts-Strogatz Model for Directed and Undirected Small-World Networks*, Phys. Rev. E **90,** 062801 (2014).

[33] R. Cohen, D. Ben-Avraham, and S. Havlin, *Percolation Critical Exponents in Scale-Free Networks*, Phys. Rev. E **66,** 036113 (2002).

[34] N. Schwartz, R. Cohen, D. Ben-Avraham, A.-L. Barabási, and S. Havlin, *Percolation in Directed Scale-Free Networks*, Phys. Rev. E **66,** 015104(R) (2002).

[35] C. Moore and M. E. Newman, *Exact Solution of Site and Bond Percolation on Small-World Networks*, Phys. Rev. E **62,** 7059 (2000).

[36] M. A. Muñoz, R. Juhász, C. Castellano, and G. Ódor, *Griffiths Phases on Complex Networks*, Phys. Rev. Lett. **105,** 128701 (2010).

[37] 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).

[38] K. Christensen and N. R. Moloney, *Complexity and Criticality* (World Scientific, Singapore, 2005).

[39] W. Cota, G. Ódor, and S. C. Ferreira, *Griffiths Phases in Infinite-Dimensional, Non-Hierarchical Modular Networks*, Sci. Rep. **8,** 9144 (2018).

[40] D. S. Callaway, M. E. Newman, S. H. Strogatz, and D. J. Watts, *Network Robustness and Fragility: Percolation on Random Graphs*, Phys. Rev. Lett. **85,** 5468 (2000).

[41] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson, *The Unreasonable Effectiveness of Tree-Based Theory for Networks with Clustering*, Phys. Rev. E **83,** 036112 (2011).

[42] R. V. Williams-García, M. Moore, J. M. Beggs, and G. Ortiz, *Quasicritical Brain Dynamics on a Nonequilibrium Widom Line*, Phys. Rev. E **90,** 062714 (2014).

[43] P. Bak, C. Tang, and K. Wiesenfeld, *Self-Organized Criticality*, Phys. Rev. A **38,** 364 (1988).

[44] E. Bienenstock, *A Model of Neocortex*, Netw.: Comput. Neural Syst. **6,** 179 (1995).

[45] P. Massobrio, V. Pasquale, and S. Martinoia, *Self-Organized Criticality in Cortical Assemblies Occurs in Concurrent Scale-Free and Small-World Networks*, Sci. Rep. **5,** 10578 (2015).

[46] C. Tetzlaff, S. Okujeni, U. Egert, F. Wörgötter, and M. Butz, *Self-Organized Criticality in Developing Neuronal Networks*, PLOS Comput. Biol. **6,** e1001013 (2010).

[47] F. Lombardi, H. Herrmann, C. Perrone-Capano, D. Plenz, and L. de Arcangelis, *Balance between Excitation and Inhibition Controls the Temporal Organization of Neuronal Avalanches*, Phys. Rev. Lett. **108,** 228703 (2012).

[48] D. Millman, S. Mihalas, A. Kirkwood, and E. Niebur, *Self-Organized Criticality Occurs in Non-Conservative Neuronal Networks During "Up" States*, Nat. Phys. **6,** 801 (2010).

[49] S. Scarpetta, I. Apicella, L. Minati, and A. de Candia, *Hysteresis, Neural Avalanches, and Critical Behavior Near a First-Order Transition of a Spiking Neural Network*, Phys. Rev. E **97,** 062305 (2018).

[50] A. Levina, J. M. Herrmann, and T. Geisel, *Dynamical Synapses Causing Self-Organized Criticality in Neural Networks*, Nat. Phys. **3,** 857 (2007).

[51] B. D. Papa, V. Priesemann, and J. Triesch, *Criticality Meets Learning: Criticality Signatures in a Self-Organizing Recurrent Neural Network*, PLOS One **12,** e0178683 (2017).

[52] V. Priesemann, M. Valderrama, M. Wibral, and M. L. V. Quyen, *Neuronal Avalanches Differ from Wakefulness to Deep Sleep–Evidence from Intracranial Depth Recordings in Humans*, PLOS Comput. Biol. **9,** e1002985 (2013).

[53] V. Priesemann, M. Wibral, M. Valderrama, R. Pröpper, M. Le Van Quyen, T. Geisel, J. Triesch, D. Nikolić, and M. H. Munk, *Spike Avalanches In Vivo Suggest a Driven, Slightly Subcritical Brain State*, Front. Syst. Neurosci. **8,** 108 (2014).

[54] A. Levina and V. Priesemann, *Subsampling Scaling*, Nat. Commun. **8,** 15140 (2017).

[55] M. Nonnenmacher, C. Behrens, P. Berens, M. Bethge, and J. H. Macke, *Signatures of Criticality Arise from Random Subsampling in Simple Population Models*, PLOS Comput. Biol. **13,** e1005718 (2017).

[56] G. Scott, E. D. Fagerholm, H. Mutoh, R. Leech, D. J. Sharp, W. L. Shew, and T. Knöpfel, *Voltage Imaging of Waking Mouse Cortex Reveals Emergence of Critical Neuronal Dynamics*, J. Neurosci. **34,** 16611 (2014).

[57] J. Touboul and A. Destexhe, *Power-Law Statistics and Universal Scaling in the Absence of Criticality*, Phys. Rev. E **95,** 012413 (2017).

[58] V. Priesemann and O. Shriki, *Can a Time Varying External Drive Give Rise to Apparent Criticality in Neural Systems?*, PLOS Comput. Biol. **14,** e1006081 (2018).

[59] M. Martinello, J. Hidalgo, A. Maritan, S. Di Santo, D. Plenz, and M. A. Muñoz, *Neutral Theory and Scale-Free Neural Dynamics*, Phys. Rev. X **7,** 041071 (2017).

[60] P. Villegas, S. di Santo, R. Burioni, and M. A. Muñoz, *Time-Series Thresholding and the Definition of Avalanche Size*, Phys. Rev. E **100,** 012133 (2019).

[61] M. Girardi-Schappo and M. H. R. Tragtenberg, *Measuring Neuronal Avalanches in Disordered Systems with Absorbing States*, Phys. Rev. E **97,** 042415 (2018).

[62] T. Bolt, M. L. Anderson, and L. Q. Uddin, *Beyond the Evoked/Intrinsic Neural Process Dichotomy*, Netw. Neurosci. **2,** 1 (2018).

[63] J. Zierenberg, J. Wilting, and V. Priesemann, *Homeostatic Plasticity and External Input Shape Neural Network Dynamics*, Phys. Rev. X **8,** 031018 (2018).

[64] R. Darshan, W. E. Wood, S. Peters, A. Leblois, and D. Hansel, *A Canonical Neural Mechanism for Behavioral Variability*, Nat. Commun. **8,** 15415 (2017).

[65] J. Zierenberg, J. Wilting, V. Priesemann, and A. Levina, *Tailored Ensembles of Neural Networks Optimize Sensitivity to Stimulus Statistics*, Phys. Rev. Research **2,** 013115 (2020).

[66] H. J. Jensen, *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems* (Cambridge University Press, Cambridge, England, 1998), Vol. 10.

[67] G. Pruessner, *Self-Organised Criticality: Theory, Models and Characterisation* (Cambridge University Press, Cambridge, England, 2012).

[68] P. Bak, *How Nature Works: The Science of Self-Organized Criticality* (Springer Science & Business Media, New York, 2013).

[69] Á. Corral and M. Paczuski, *Avalanche Merging and Continuous Flow in a Sandpile Model*, Phys. Rev. Lett. **83,** 572 (1999).

[70] Z.-L. Hu, J.-G. Liu, G.-Y. Yang, and Z.-M. Ren, *Effects of the Distance Among Multiple Spreaders on the Spreading*, Europhys. Lett. **106,** 18002 (2014).

[71] J. C. Miller, *Epidemics on Networks with Large Initial Conditions or Changing Structure*, PLOS One **9,** e101421 (2014).

[72] M. Zheng, C. Wang, J. Zhou, M. Zhao, S. Guan, Y. Zou, and Z. Liu, *Non-Periodic Outbreaks of Recurrent Epidemics and Its Network Modelling*, Sci. Rep. **5,** 16010 (2015).

[73] T. Hasegawa and K. Nemoto, *Outbreaks in Susceptible-Infected-Removed Epidemics with Multiple Seeds*, Phys. Rev. E **93,** 032324 (2016).

[74] W. Choi, D. Lee, and B. Kahng, *Critical Behavior of a Two-Step Contagion Model with Multiple Seeds*, Phys. Rev. E **95,** 062115 (2017).

[75] T. Hasegawa and K. Nemoto, *Sudden Spreading of Infections in an Epidemic Model with a Finite Seed Fraction*, Eur. Phys. J. B **91,** 58 (2018).

[76] M. Di Muro, L. Alvarez-Zuzek, S. Havlin, and L. Braunstein, *Multiple Outbreaks in Epidemic Spreading with Local Vaccination and Limited Vaccines*, New J. Phys. **20,** 083025 (2018).

[77] E. Cator and P. Van Mieghem, *Susceptible-Infected-Susceptible Epidemics on the Complete Graph and the Star Graph: Exact Analysis*, Phys. Rev. E **87,** 012811 (2013).

[78] J. Zhang, J. M. Moura, and J. Zhang, *Contact Process with Exogenous Infection and the Scaled SIS Process*, J. Complex Netw. **5,** 712 (2017).

[79] A. J. Rodriguez-Morales, A. C. Bandeira, and C. Franco-Paredes, *The Expanding Spectrum of Modes of Transmission of Zika Virus: A Global Concern*, Ann. Clinical Microbiol. Antimicrobials **15,** 13 (2016).

[80] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, *The Physics of Spreading Processes in Multilayer Networks*, Nat. Phys. **12,** 901 (2016).

[81] J. Sanz, C.-Y. Xia, S. Meloni, and Y. Moreno, *Dynamics of Interacting Diseases*, Phys. Rev. X **4,** 041005 (2014).

[82] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, *Identification of Influential Spreaders in Complex Networks*, Nat. Phys. **6,** 888 (2010).

[83] Y. Hu, S. Ji, Y. Jin, L. Feng, H. E. Stanley, and S. Havlin, *Local Structure Can Identify and Quantify Influential Global Spreaders in Large Scale Social Networks*, Proc. Natl. Acad. Sci. U.S.A. **115,** 7468 (2018).

[84] J. Nagler, A. Levina, and M. Timme, *Impact of Single Links in Competitive Percolation*, Nat. Phys. **7,** 265 (2011).

[85] P. Grassberger, C. Christensen, G. Bizhani, S.-W. Son, and M. Paczuski, *Explosive Percolation Is Continuous, but with Unusual Finite Size Behavior*, Phys. Rev. Lett. **106,** 225701 (2011).

[86] R. M. D'Souza and J. Nagler, *Anomalous Critical and Supercritical Phenomena in Explosive Percolation*, Nat. Phys. **11,** 531 (2015).

[87] F. Morone, K. Roth, B. Min, H. E. Stanley, and H. A. Makse, *Model of Brain Activation Predicts the Neural Collective Influence Map of the Brain*, Proc. Natl. Acad. Sci. U.S.A. **114,** 3849 (2017).

[88] N. Chen and M. Olvera-Cravioto, *Directed Random Graphs with Given Degree Distributions*, Stoch. Syst. **3,** 147 (2013).

[89] Here, we use $\kappa = 0.5$ and $\delta = 0.05$, to be consistent with the notation of Chen *et al.* [88].

[90] D. R. D'Souza, T. Tomé, and R. M. Ziff, *A New Scale-Invariant Ratio and Finite-Size Scaling for the Stochastic Susceptible–Infected–Recovered Model*, J. Stat. Mech. (2011) P03006.

[91] G. Evans and J. Webster, *A Comparison of Some Methods for the Evaluation of Highly Oscillatory Integrals*, J. Comput. Appl. Math. **112,** 55 (1999).

[92] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science* (Addison-Wesley, New York, 1989).

[93] Y. L. Luke, *Special Functions and Their Approximations* (Academic Press, New York, 1969), Vol. 1.

[94] G. N. Watson, *Asymptotic Expansions of Hypergeometric Functions*, Trans. Cambridge Philos. Soc. **22,** 277 (1918).

Correction: The email address in the byline footnotes for the last author contained an error and has been fixed.