Critical excitation/inhibition balance in dense neural networks

Lorenz Baumgarten* and Stefan Bornholdt†

Institut für Theoretische Physik, Universität Bremen, 28759 Bremen, Germany
(Dated: April 1, 2019)

The “edge of chaos” phase transition in artificial neural networks is of renewed interest in the light of recent evidence for criticality in brain dynamics. Statistical mechanics traditionally studied this transition with connectivity $k$ as control parameter and an exactly balanced excitation/inhibition ratio. While critical connectivity has been found to be low in these model systems, typically around $k = 2$, which is unrealistic for natural systems, a recent study utilizing the excitation/inhibition ratio as control parameter found a new, nearly degree-independent critical point when connectivity is large. However, the new phase transition is accompanied by an unnaturally high level of activity in the network.

Here we study random neural networks with the additional properties of (i) a high clustering coefficient and (ii) neurons that are solely either excitatory or inhibitory, a prominent property of natural neurons. As a result we observe an additional critical point for networks with large connectivity, regardless of degree distribution, which exhibits low activity levels that compare well with neuronal brain networks.

Between the ordered and chaotic regimes of threshold neural networks lies the “edge of chaos”, a critical point where the length and size distributions of activity avalanches are governed by characteristic power laws. This dynamical phase transition has been thoroughly studied in random neural networks [1–4], non-symmetric spin glasses [5], and random Boolean networks [6–10]. Traditionally, threshold networks have been studied with precisely balanced excitation and inhibition, usually by randomly assigning activating and inhibiting links with equal probabilities. In these networks, criticality occurs for small average degrees $k$ [1]. However, when allowing the fraction of excitatory links $F_+$ as a second control parameter of the phase transition, it was recently discovered that there exist two critical lines in the $k-F_+$-plane: one almost parallel to the $F_+$ axis at low $k$ and one almost independent of $k$ at some $F_+ > 0.5$ [11], see FIG. 1.

This new critical point’s relevance becomes apparent in the context of neural brain networks which exhibit a high average degree ($k \approx 10^4$ in human brains [12]) and a characteristic imbalance between excitation and inhibition (20–30% of neurons are inhibitory in monkey brains [13]). There is a large amount of evidence suggesting that the brain operates near a critical point, namely avalanche sizes and durations governed by power laws [14–18], the possibility of tuning from a subcritical regime through the critical point to a supercritical regime [19], mathematical relations between critical exponents, and collapsable avalanche shapes [15, 20]. Further, Fraiman et al. showed striking similarities between correlation networks extracted from brains and the Ising model at the critical point [21]. The interest in the role of criticality in the brain is illustrated by a large amount of research devoted to criticality in network models inspired by biological networks [22–29].

Unfortunately, this high-degree critical point exists in a high-activity regime which is unrealistic for brain networks. We find, however, an additional critical point that persists at low activities, when including additional network properties characteristic of brain networks, thereby providing a more likely network model candidate for

---

* lbaumgarten@itp.uni-bremen.de
† bornholdt@itp.uni-bremen.de

FIG. 1. Sensitivity as a function of fraction of excitatory links $F_+$ and connectivity $k$, similar to figure 1 C in [11], for threshold $h = 0$ and $N = 10^3$ nodes. Lines show a comparison of equation (3) (green solid line) and the numerical results for equation (2) (red dashed line) with the simulation results. Both lines approximate the simulation’s critical line well for large $k$. 
scribing the processes behind brain criticality.

We use threshold networks consisting of $N$ nodes connected by $kN$ directed edges, whose node-states are updated in parallel according to

$$\sigma_i(t+1) = \begin{cases} 1, & \text{if } \sum_{j=1}^{N} w_{ij} \sigma_j(t) > h \\ 0, & \text{if } \sum_{j=1}^{N} w_{ij} \sigma_j(t) \leq h, \end{cases}$$

(1)

where $\sigma_i(t)$ is the node $i$’s state at time $t$ and $w_{ij}$ is the weight of the connection from node $j$ to node $i$. The weights $w_{ij}$ can be 0 if there is no connection between nodes $i$ and $j$, or $\pm 1$ otherwise. The weights of existing connections are chosen randomly with excitatory links $w_{ij} = +1$ chosen with probability $F_+$. Initial states of the nodes are chosen according to a random initial activity $A_0 = \frac{1}{N} \sum_i \sigma_i$.

A simple quantity we use to measure criticality is the sensitivity $\lambda$ [30, 31]. Imagine switching one node’s state in the current timestep, then $\lambda$ is defined as the average number of nodes whose states will then be different in the next timestep from what they would have been otherwise. If sensitivity is smaller or larger than 1, perturbations will quickly die out or dominate the entire network, respectively. Hence, at $\lambda = 1$ the network is in a critical state.

First, in order to establish whether the vertical white line defined by $\lambda = 1$ seen in FIG. 1 indeed is a critical point, we measure the averages of multiple quantities of interest, as well as the average sensitivity for $10^3$ timesteps after letting the network relax from its initial condition within $2 \cdot 10^3$ timesteps (tests show that increasing this time or waiting until an attractor is reached — where possible, attractors cannot be found in a reasonable amount of time for $\lambda \gg 1$ — does not change the results) for different values of $F_+$. Afterwards, we can plot the measured quantities as a function of the sensitivity. The measured quantities are: The network’s activity $A$, the fraction of nodes which do not change their state within the $10^3$ timesteps $N_A$, and the average number of state-changes per node and timestep $F/Nt$. This measurement is shown in FIG. 2.

For $\lambda < 1$, essentially all nodes are static (i.e. remaining in one state, either active or inactive) and almost no flips happen, whereas for $\lambda > 1$ the number of static nodes drops and the number of flips increases, so $\lambda = 1$ is a boundary between order and chaos. Also note that the network’s activity is very high at the critical point. It seems, therefore, that this critical point cannot underlie a mechanism that defines criticality in the brain, as almost all neurons constantly firing is not realistic.

Further, we measure avalanche sizes and durations at the critical point, as described in the supplemental material [32], see FIG. 3. We observe power laws in both avalanche size and duration distributions.

Finally, we measure the attractor and transient lengths, as well as the average sensitivity within the attractor for a number of different network realizations for fixed parameters. We only use parameter and attractor lengths of networks whose average sensitivity $\lambda$ is within $1 - \delta \leq \lambda \leq 1 + \delta$ with $\delta = 0.05$. Attractor and transient length distributions are shown in FIG. 3. Both the attractor lengths as well as the right flank of the transient length distributions show clear power laws, as is to be expected for critical networks [33].

All the above discussed properties lead us to conclude that this is indeed a critical point. We use Derrida’s annealed approximation [6], adopted for threshold networks [2], to estimate the critical $F_+$ as a function of $k$, and arrive at the equation

$$\frac{1}{k} = \left( \frac{k}{2} \right)^{k+h+1} \frac{k+h+1}{2} \left( 1 - F_+ \right)^{\frac{k-h-1}{2} k + h + 1}.$$

(2)

Under the assumption of large average degree $k \gg h$,
$k \gg 1$, this can be simplified to

$$F_+ = \frac{1}{2} \left[ 1 + \left\{ 1 - \left( \frac{2\pi}{k} \right)^{\frac{3}{2}} \right\}^{\frac{1}{2}} \right].$$  \hspace{1cm} (3)$$

See supplemental material [32] for details. Figure 1 shows a comparison of equation (3), as well as the numerical solution of equation (2), with our simulation results.

In FIG. 1 all activity dies out on the left flank, preventing the average sensitivity from crossing through one, except for entirely inactive networks whose sensitivity is not shown. However, as a central observation of our study, we find that networks can be kept from dying out for low $F_+$ by introducing two properties to the network: increasing the networks’ clustering coefficient $C$ and requiring that nodes have either only excitatory or only inhibitory outgoing edges. Both of these properties are prominent features of brain networks [34–38]. Networks with only a high clustering coefficient, without the second property, can also show surviving activity on the left flank for some initial configurations and for exceedingly high clustering coefficients and thresholds, but even then the left flank drops sharply towards zero. In the following, let us denote networks with excitatory or inhibitory nodes as EIN networks, as opposed to networks with excitatory/inhibitory edges (EIE networks). Since the network’s activity does not abruptly die out on the left flank anymore for such networks, a second critical point can be found here, as shown in FIG. 4 A. Note that, in contrast to the first critical point, this second critical point exists in a low-activity state, making it more interesting for real life applications, such as studying mechanisms underlying brain criticality.

To construct networks with high clustering coefficients, we here use directed Watts-Strogatz (WS) networks [39, 40]. By manipulating the rewiring probability $\beta$, we can vary a network’s clustering coefficient and average path length. The Watts-Strogatz model’s strength is that when varying $\beta$, there is a region in which the clustering coefficient is nearly constant while the average path length changes drastically and a second region in which the clustering coefficient changes and the average path length is nearly constant, enabling us to isolate these two parameters’ effects. We find that the second critical point comes into existence in the region in which the clustering coefficient changes, while it is unaffected by changes within the region in which the clustering coefficient is constant. Therefore, a high clustering coefficient is sufficient to enable the second critical point’s existence. The influence of thresholds and clustering coefficients, as well as the difference between EIE and EIN networks is shown in FIG. 4 B and 5.

So far, our networks had degree distributions centered around an average value; however, random or Watts-Strogatz models rarely describe real-life networks. Scale-free or similar networks are significantly more abundant in nature. In fact, for neuronal networks, cumulative degree distributions ranging from power laws [41–43] over exponentially truncated power laws [44–48] to exponential laws [49–52] have been found, with the observation that distributions following exponentially truncated power laws increasingly resemble true power laws for measurements on finer scales [44].

In analogy to the brain, we focus on EIN networks with a broad link distribution, and here create them by the algorithm described in the supplemental material [32], which is based on the algorithm described in [53] to create networks that are better suited to our purposes than simple Barabási-Albert [54, 55], or similar [56], networks. We find that for scale-free graphs the right critical point still exists, see FIG. 6, and that the sensitivity $\lambda$ splits into two paths on the right flank and is therefore no longer solely dependent on $F_+$. The two different paths are dependent on whether the network’s largest node is excitatory or inhibitory (in our algorithm there is a clear hierarchy between nodes, dictated by when they were introduced to the network, and therefore the first node is always clearly larger than the rest, so no multiple nodes are competing for the spot of largest node). Similarly to the existence of the left flank in WS networks, this split in the sensitivity is amplified by high clustering coefficients and thresholds.

Figure 6 also shows the existence of the left flank’s second critical point for high clustering coefficients $C$ and thresholds $h$, see FIG. 6. For low clustering coefficients, the left flank still dies out. High clustering coefficients and thresholds lower the sensitivity curve’s slope, so that for certain parameters the sensitivity, and therefore criticality, is almost constant over a wide area of $F_+$, see FIG.
FIG. 5. Sensitivity $\lambda$ and activity $A$ for different network configurations. White denotes a critical sensitivity. A: For EIE networks—except for very high clustering coefficients and thresholds—the left flank dies out before reaching the critical point. B: Switching to an EIN network stabilizes the left flank; however, it still collapses for high average degrees $k$ without a high clustering coefficient. Some white artifacts can be seen because the left flank does not die out within $2 \cdot 10^3$ timesteps; it does, however, die out after a larger number of timesteps, and therefore no second critical point exists here, see supplemental material [32] for more information. C: A higher clustering coefficient $C = 0.7$ stabilizes the left flank even for higher average degrees $k$. D: With a higher threshold $h = 10$ even a lower clustering coefficient $C = 0.25$ can have a stable left flank. The distance between the critical points shrinks for higher thresholds and both critical points are also moved to higher $F^+$. E: For EIN networks, a higher clustering coefficient ($C = 0.5$) lowers the network’s average sensitivity, leading this configuration to only barely pass above $\lambda = 1$ between the critical points. From the shape of the left critical line from B to D, it can also be seen that the left critical line is merely a continuation of the horizontal line from FIG. 1 folded upwards.

FIG. 6. A–C: Sensitivity $\lambda$ and D–F: Activity $A$ as a function of $F^+$ for an exponentially truncated power law network with low clustering coefficient (red) and a scale-free network with high clustering coefficient, where the largest node is excitatory (green) or inhibitory (blue) at $k = 40$, $N = 10^4$, and A,D: $h = 1$, B,E: $h = 7$, C,F: $h = 10$. The sensitivity and activity for the highly clustered network were measured as the average within the network’s attractor.

To summarize, in threshold neural networks, a previously unknown phase transition between a chaotic and a quiescent regime has been found for highly clustered networks with exclusively excitatory/inhibitory nodes. This critical point exhibits a persisting, yet low level of average activity (which in unclustered networks would die out). Besides the requirement of a certain level of clustering, it is robust both for random as well as broad (scale-free) degree distributions. This new critical point is of particular interest to neuroscience because it is relatively independent of the degree $k$ and may, therefore, occur at the large average degree present in brains. Furthermore, the main prerequisites for this critical point’s existence are present in the brain: A highly clustered architecture and nodes being either exclusively excitatory or inhibitory (Dale’s principle).

It can only be speculated what role criticality may play in nature. It has been discussed that it could optimize a network’s information processing capabilities. Yet also, dynamical phase transitions are a simple means that physics provides, allowing a complex system to tune to an intermediate activity regime with great ease.

Last but not least, research has shown that the balance between excitation and inhibition in the brain, which needs to be a specific value for a network to be critical in our model, is vital for a functioning brain [57–61] and that disturbing this balance can negatively impact information processing [62]. Interestingly, the ratio of excitatory and inhibitory neurons in brain networks is ob-
served to be almost constant throughout an organism’s development, and feedback algorithms that regulate this ratio are currently discussed [63]. We hope that the critical point described in this paper may help our further understanding of the mechanisms that regulate the excitation/inhibition balance in the brain.

[1] K. Kürten, Phys. Lett. A 129, 157 (1988).
[2] S. Bornholdt and T. Rohlf, Physica A 310, 245 (2002).
[3] A. Szejka, T. Mihaljev, and B. Drossel, New J. Phys. 10, 063009 (2008).
[4] T. Rohlf, Phys. Rev. E 78, 066118 (2008).
[5] B. Derrida, J. Phys. A 20, L721 (1987).
[6] B. Derrida and Y. Pomeau, Europhys. Lett. 1, 45 (1986).
[7] K. E. Kürten, J. Phys. A 21, L615 (1988).
[8] M. Aldana, S. Coppersmith, and L. Kadanoff, in Perspectives and Problems in Nonlinear Science, edited by E. Kaplan, J. Marsden, and K. Sreenivasan (Springer, 2003).
[9] B. Drossel, in Reviews of Nonlinear Dynamics and Complexity, Vol. 1, edited by H. Schuster (Wiley-VCH, 2008).
[10] S. Bornholdt and S. Kauffman, Journal of Theoretical Biology 467, 15 (2019).
[11] S. Bornholdt, J. Neto, M. de Aguiar, and J. Brum, arXiv 1712.08816 (2017).
[12] P. R. Huttenlocher, Brain Res. 163, 195 (1979).
[13] S. Hendry, H. D. Schwark, E. Jones, and J. Yan, J. Neurosci. 7, 1503 (1987).
[14] J. Beggs and D. Plenz, J. Neurosci. 23, 11167 (2003).
[15] N. Friedman, S. Ito, B. Brinkman, M. Shimono, R. DeVille, K. Dahmen, J. Beggs, and T. Butler, Phys. Rev. Lett. 108, 208102 (2012).
[16] V. Priesemann, M. Wibral, M. Valderrama, R. Pröpper, M. Quyen, T. Geisel, J. Triesch, D. Nikolić, and M. Munk, Front. Syst. Neurosci. 8 (2014).
[17] N. Timme, N. Marshall, N. Bennett, M. Ripp, E. Lautzenhiser, and J. Beggs, Front. Physiol. 7, 425 (2016).
[18] M. Yaghoubi, T. de Graaf, J. Orlandi, F. Girotto, M. Colcicos, and J. Davidson, Sci. Rep. 8, 3417 (2018).
[19] C. Haldeman and J. Beggs, Phys. Rev. Lett. 94, 058101 (2005).
[20] A. Shankat and J. Thivierge, Front. Comput. Neurosci. 10, 29 (2016).
[21] D. Fraiman, P. Balenzuela, J. Foss, and D. Chialvo, Phys. Rev. E 79, 061922 (2009).
[22] T. Gross and B. Blasius, J. Royal Soc. Interface 5 20, 259 (2008).
[23] X. Li, Q. Chen, and F. Xue, Philos. Trans. A Math. Phys. Eng. Sci. 375 (2017).
[24] W. Clawson, N. Wright, R. Wessel, and W. Shew, PLoS Comput. Biol. 13, e1005574 (2017).
[25] L. Brochini, A. de Andrade Costa, M. Abadi, A. Roque, J. Stolfi, and K. Kinouchi, Sci. Rep. 6 (2016).
[26] S. Gautam, T. Hoang, K. McClanahan, S. Grady, and W. Shew, PLoS Comput. Biol. 11, e1004576 (2015).
[27] O. Shriki and D. Yellin, PLoS Comput. Biol. 12, e1004698 (2016).
[28] B. V. Rodriguez, A. Avena-Koenigsberger, O. Sporns, A. Griffa, P. Hagmann, and H. Larralde, Sci. Rep. 7, 13020 (2017).
[29] M. Ferraz, H. Melo-Silva, and A. Kihara, PLoS One 12, e0184367 (2017).
[30] B. Luque and R. Solé, Phys. Rev. E 55, 257 (1997).
[31] I. Shmulevich and S. Kauffman, Phys. Rev. Lett. 93, 048701 (2004).
[32] See Supplemental Material at [URL] for details on the annealed approximation and simulation methods.
[33] A. Bhattacharjya and S. Liang, Phys. Rev. Lett. 77, 1644 (1996).
[34] K. E. Stephan, C.-C. Hilgetag, G. A. P. C. Burns, M. A. O’Neill, M. P. Young, and R. Köttér, Philos Trans R Soc Lond B Biol Sci. 355, 111 (2000).
[35] O. Sporns and J. D. Zwi, Neuroinf. 2, 145 (2004).
[36] O. Sporns, C. J. Honey, and R. Köttér, PLoS One 2, e1049 (2007).
[37] D. S. Bassett and E. Bullmore, The Neuroscientist 12, 512 (2006).
[38] H. Dale, Proc. R. Soc. Med. 28, 319 (1935).
[39] D. Watts and S. Strogatz, Nature 393, 440 (1998).
[40] H. Song and X. Wang, Phys. Rev. E 90, 062801 (2014).
[41] L. Varshney, B. Chen, E. Panagiau, D. Hall, and D. Chklovskii, PLoS Comput. Biol. 7, e1001066 (2011).
[42] M. van den Heuvel, C. Stam, M. Boersma, and H. H. Pol, Neuroimage 43, 528 (2008).
[43] V. Eguíluz, D. Chialvo, G. Cecchi, M. Baliki, and A. Apkarian, Phys. Rev. Lett. 9, 018102 (2005).
[44] S. Hayasaka and P. Laurienti, Neuroimage 50, 499 (2010).
[45] Y. He, Z. Chen, and A. Evans, Cereb. Cortex 17, 2407 (2007).
[46] Y. Iturria-Medina, R. Sotero, E. Canales-Rodríguez, Y. Alemán-Gómez, and L. Melía-Garcia, Neuroimage 40, 1064 (2008).
[47] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, Neurosci. 26, 63 (2006).
[48] G. Gong, Y. He, L. Concha, C. Lebel, D. Gross, A. Evans, and C. Beaulieu, Cereb. Cortex 19, 524 (2009).
[49] D. Modha and R. Singh, Proc. Natl. Acad. Sci. USA 107, 13485 (2010).
[50] L. Amaral, A. Scala, M. Barthélémy, and H. Stanley, Proc. Natl. Acad. Sci. USA 97, 11149 (2000).
[51] P. Hagmann, L. Cammoun, X. Gagandet, R. Meuli, C. Honey, V. Wedeen, and O. Sporns, PLoS Biol. 6, e159 (2008).
[52] D. de Santos-Sierra, I. S. n a Nadal, I. Levyra, J. A. Almendral, S. Anava, A. Ayali, D. Papo, and S. Boccaletti, PLoS One 9, e85828 (2014).
[53] Y. Lo, C. Li, and S. Lin, Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom) (2012).
[54] A. Barabási and R. Albert, Science 286, 509 (1999).
[55] M. Newman, SIAM Review 45, 167 (2003).
[56] D. de S. Price, J. Amer. Soc. Inform. Sci. 27, 292 (1976).
[57] B. Haider, A. Duque, A. R. Hasenstaub, and D. A. McCormick, J. Neurosci. 26, 4535 (2006).
[58] J. L. R. Rubenstein and M. M. Merzenich, Genes Brain Behav. 2, 255 (2003).
[59] E. C. Davenport, B. R. Szulc, J. Drew, J. Taylor, T. Morgan, N. F. Higgs, G. López-Doménech, and J. T. Kittler, Cell Rep. 26, 2037 (2019).

[60] M. W. Antoine, T. Langberg, P. Schnepel, and D. E. Feldman, Neuron 101, 648 (2019).

[61] J. D. Murray and X.-J. Wang, in Computational Psychiatry, edited by A. Anticevic and J. D. Murray (Academic Press, 2018) pp. 3 – 25.

[62] O. Yizhar, L. E. Fenno, M. Prigge, F. Schneider, T. J. Davidson, D. J. O’Shea, V. S. Sohal, I. Goshen, J. Finkelstein, J. T. Paz, K. Stehfest, R. Fudim, C. Ramakrishnan, J. R. Huguenard, P. Hegemann, and K. Deisseroth, Nature 477, 171 (2011).

[63] S. Sahara, Y. Yanagawa, D. D. M. O’Leary, and C. F. Stevens, J. Neurosci. 32, 4755 (2012).
Supplemental Material to: Critical excitation/inhibition balance in dense neural networks

Lorenz Baumgarten* and Stefan Bornholdt†

Institut für Theoretische Physik, Universität Bremen, 28759 Bremen, Germany

(Dated: April 1, 2019)

I. ANNEALED APPROXIMATION

We assume that all nodes are active, which is close to true at the critical point, and that the degree distribution is narrow enough that it can be approximated by a single peak at \( k \). Consider a single node with a fixed number \( k \) of incoming signals. The probability of any specific signal \( Z \) is

\[
p(Z) = \binom{k}{k_+} F_+^{k_+} F_-^{k_-},
\]

where \( F_- = 1 - F_+ \) and \( k_+ \) and \( k_- \) are the number of incoming excitatory and inhibitory signals, respectively. We want to calculate the probability \( p_s \) that the node will change its state in the next timestep if one of the incoming signals is turned off. This can only happen if the sum of incoming signals is

\[
k_+ - k_- = h
\]

or \( k_+ - k_- = h + 1 \).

Note that for given \( k \) and \( h \) only one of these two cases can occur since \( k \) and \( h \) need to both be odd or both be even for the first case and have different parities for the second case.

In the first case, an inhibitory incoming signal needs to be turned off to effect a flip. Therefore, the fraction of connections whose disabling would produce a flip is

\[
\frac{k_-}{k} = \frac{k-h}{2k},
\]

and for the second case, where an excitatory connection has to be turned off to effect a flip, the fraction is

\[
\frac{k_+}{k} = \frac{k+h+1}{2k}.
\]

The respective damage spreading probabilities are

\[
p_s(-) = \frac{\sum_{Z \in Z_-} p(Z) k - h}{\sum_Z p(Z)},
\]

\[
= \left( \binom{k}{k+h} F_+^{k+h} F_-^{k-h} \right) \frac{k-h}{2k},
\]

\[
p_s(+) = \frac{\sum_{Z \in Z_+} p(Z) k + h + 1}{\sum_Z p(Z)},
\]

\[
= \left( \binom{k}{k+h+1} F_+^{k+h+1} F_-^{k-h-1} \right) \frac{k+h+1}{2k}.
\]

Because we assume high activity, which requires \( k_+ - k_- > h \), equation 7 is used in the main text. For \( k \gg h \) and \( k \gg 1 \), the probabilities can be approximated by

\[
p_s(-) \approx p_s(+) \approx p_s = \frac{1}{2} \left( \binom{k}{k/2} (F_+ F_-)^{k/2} \right).
\]

A sensitivity of \( \lambda = 1 \) is equivalent to a damage-spreading probability of

\[
p_s = \frac{1}{k}.
\]

We use Stirling’s approximation on the binomial coefficient in \( p_s \) to arrive at

\[
\frac{1}{k} = \frac{2^k}{k} \sqrt{\frac{k}{2\pi}} \left( F_+ F_- \right)^{k/2}
\]

\[
\Leftrightarrow 0 = F_-^2 - F_+ + \frac{1}{4} \left( \frac{2\pi}{k} \right)^{1/2}
\]

\[
\Rightarrow F_+ = \frac{1}{2} \left[ 1 \pm \sqrt{1 - \left( \frac{2\pi}{k} \right)^{1/2}} \right]^{1/2}.
\]

Only the solution with \( F_+ > 0.5 \) is realistic, since the assumption of high activity does not hold for \( F_+ < 0.5 \).

II. METHODS

A. measuring avalanche sizes and lengths

We initialize random networks with random activity and let their dynamics evolve until the network runs into an attractor. If the average sensitivity within the attractor is within a small margin of one \( (1 - \delta < \lambda < 1 + \delta \)
with $\delta = 10^{-2}$, we change the state of one node and again simulate the network dynamics until the network either returns to its old attractor or reaches a new one. During this, we measure the number of nodes that had a different state than they would have had in the untouched initial attractor at the corresponding time, and the number of timesteps it takes to arrive at an attractor. This is repeated for all nodes in the network.

B. existence of the second critical point

For large clustering coefficients and high thresholds, transient lengths can become exceedingly long on the left flank, and therefore the assertion in the main text that letting networks evolve for $2 \cdot 10^3$ timesteps before measuring is sufficient to gauge the network’s dynamics may not always be correct here. Some left flanks may seemingly exist, but will eventually die out after a long time. Although the measurements shown in the main text were, unless otherwise stated, not explicitly done after the transient period, because waiting until the network reaches an attractor is not feasible for large sensitivities, the existence of all critical points shown was verified by letting the network reach an attractor and measuring the average sensitivity within the attractor.

Further, we also tested whether this left flank was merely an aberration caused by the WS model’s unrealistic ring structure by repeating our measurements for a two-dimensional network in which node position’s were randomly chosen within a square area and node connections were established according to a decreasing exponential probability function of node distance (with periodic boundary conditions). By varying this probability function, different clustering coefficients could be achieved. We verified that the left flank, and therefore also the second critical point, also exist in these more realistic networks.

C. scale-free and exponentially truncated scale-free graphs

Many conventional algorithms to produce scale-free graphs, such as the Barabási-Albert model [1], which produces scale-free graphs, or derivatives of it which can produce the other degree distributions, are unfortunately not feasible for our types of networks. In the Barabasi-Albert model, for example, for an average degree $k$ every node will have a degree of at least $k/2$, which is acceptable for small average degrees, but for large average degrees excludes the possibility for a large amount of possible degrees, defeating the purpose of a large degree-spread in scale-free networks. Further, for the total amount of edges in our systems $Nk$, initializing scale-free graphs with most algorithms can take an unfeasibly long time. We require an algorithm that (1) can produce a scale-free graph in which low-degree nodes can exist, (2) can initialize large networks fast, (3) can produce networks with variable clustering coefficient, as we have already seen that this can have a large impact on criticality, and if possible (4) can also produce other degree distributions similar to scale-free graphs.

For this purpose, we adapt the algorithm described by Lo et al. [2] to fit our criteria. We start with a single node and add one edge between two nodes, both the edge’s origin and target chosen by preferential attachment, in every two timesteps, so that the network’s total degree increases by one per timestep. After every $2m$ timesteps, we also add one node to the system. Ignore for now that it is not always possible to add edges to the system for small $t$. One significant difference in our algorithm is that the newly added edges need not connect to the newly added node, but can instead connect any two nodes in the system, allowing low-degree nodes to exist in the final network.

We do not, however, use the nodes’ actual degrees to choose which nodes will gain an edge, but instead calculate expected degrees for the nodes. The expected degree at timestep $t_j$ of a node $v_k$ that was introduced to the network at time $t_k$ is

$$
ExpDeg(t_j, v_k) = \frac{ExpDeg(t_{j-1}, v_k) + ExpDeg(t_{j-1}, v_k)}{total(t_{j-1})}
$$

$$
= \frac{ExpDeg(t_{j-1}, v_k) + ExpDeg(t_{j-1}, v_k)}{total(t_{j-1})}
$$

$$
= \frac{ExpDeg(t_{j-1}, v_k) + ExpDeg(t_{j-1}, v_k)}{total(t_{j-1})} \prod_{i=k}^{j-1}(1 + 1/total(t_i)),
$$

where $total(t_j)$ is what we call the system’s total adjusted degree at time $t_j$. In our definition of a node’s adjusted degree (and the expected degree), we include a bias $\delta$ that is added to a node’s number of edges. The total adjusted degree is therefore

$$
total(t_j) = t_j + \delta n_j,
$$

where $n_j$ is the number of nodes in the system at time $t_j$. This bias is necessary because all nodes start with zero edges and would therefore never receive any new edges for $\delta = 0$. Using the nodes’ expected instead of their actual degrees enables us to parallelize our algorithm, significantly speeding up a network’s initialization. Note that for the expected degree the issue of being unable to add any edges for small network sizes is irrelevant. We simply pretend that nodes had the degree they could have, had we actually added an edge in every two timesteps, which still simulates preferential attachment. This simplistic approach enables us to simplify products in the following calculations which would otherwise be computationally expensive.

More useful than every single node’s degree, however, is
the cumulative degree

\[ ECumDeg(t_j, v_k) = \sum_{i=1}^{k} ExpDeg(t_j, v_i) \]

\[ = \sum_{i=1}^{k} ExpDeg(t_k, v_i) \prod_{i=k}^{j-1} (1 + 1/total(t_i)) \]

\[ = total(t_k) \prod_{i=k}^{j-1} (1 + 1/total(t_i)) \]

\[ = total(t_k) \cdot \xi(t_k, t_j), \quad (13) \]

as it allows us to do a fast binary search to find the node an edge needs to be connected to, without having to calculate and sum up expected degrees for a large number of nodes. The factor \( \xi(t_k, t_j) \) can easily be calculated when the number of nodes currently in the network \( n \) is constant

\[ \xi(t_k, t_j) = \prod_{i=k}^{j-1} \left( 1 + \frac{1}{i + \delta n} \right) \]

\[ = \left( \frac{k + \delta n + 1}{k + \delta n} \right) \left( \frac{k + \delta n + 2}{k + \delta n + 1} \right) \cdots \left( \frac{j + \delta n}{j + \delta n - 1} \right) \]

\[ = \frac{j + \delta n}{k + \delta n}. \quad (14) \]

To calculate \( \xi(t_k, t_j) \) even if nodes are added to the system between \( t_k \) and \( t_j \), we split the product into parts with constant \( n \)

\[ \xi(t_k, t_j) = \frac{k + 2m + \delta n_k}{k + \delta n_k} \times \frac{k + 4m + \delta (n_k + 1)}{k + 2m + \delta n_k} \]

\[ \times \cdots \times \frac{k + 2m(n_j - n_k) + \delta n_j}{k + 2m(n_j - n_k) - 1 + \delta(n_j - 1)} \]

\[ \times \frac{j + \delta n_j}{k + 2m(n_j - n_k) + \delta n_j}. \quad (15) \]

The product in the first two lines can be written as

\[ \prod_{i=0}^{n_j - n_k - 1} \frac{k + 2mi + 2m + \delta(n_k + i)}{k + 2mi + \delta(n_k + i)} \]

\[ = \prod_{i=0}^{n_j - n_k - 1} \left( 1 + \frac{1}{c + i(1 + \delta/2m)} \right) \]

\[ = \frac{\Gamma\left(\frac{5}{\delta}\right)\Gamma\left(b + \frac{5}{\delta} + \frac{1}{2}\right)}{\Gamma\left(\frac{5}{\delta} + \frac{1}{2}\right)\Gamma\left(b + \frac{5}{\delta}\right)}, \quad (16) \]

with \( b = n_j - n_k \), \( c = (k + \delta n_k)/2m \), \( d = 1 + \delta/2m \).

The cumulative degree is then

\[ ECumDeg(t_j, v_k) = total(t_k) \]

\[ \times \frac{\Gamma\left(\frac{5}{\delta}\right)\Gamma\left(b + \frac{5}{\delta} + \frac{1}{2}\right)}{\Gamma\left(\frac{5}{\delta} + \frac{1}{2}\right)\Gamma\left(b + \frac{5}{\delta}\right)} \times \frac{j + \delta n_j}{k + 2m(n_j - n_k) + \delta n_j}. \quad (17) \]

**FIG. 1.** Cumulative in-degree distributions for a scale-free graph with \( \delta = 1 \), \( m = 20 \), \( k = 40 \), \( N = 2.5 \cdot 10^4 \), and \( C \approx 0.52 \) (yellow) and an exponentially truncated scale-free graph with \( \delta = 40 \), \( m = 2 \), \( k = 40 \), \( N = 2.5 \cdot 10^4 \), and \( C \approx 0.005 \) (red). Solid black lines show power law and exponentially truncated power law fits, respectively. The power law's slope is -1.3.

At time \( t_j \), a node can now be chosen via preferential attachment by choosing a random number \( \eta \) between zero and one, and finding the first \( v_k \) for which \( ECumDeg(t_j, v_k)/total(t_j) \geq \eta \).

It is likely that edges between the largest nodes will be added multiple times during a network’s initialization. As it would be computationally expensive to check whether an edge already exists in the system, we ignore additional edges and remove them after initialization so weights remain as ±1. This, and also the inability to add an edge every two timesteps for low \( t \), leads to the network’s eventual average degree being unpredictable. Therefore, we repeat the initialization process, adding up all of the single initializations’ edges, until \( k \approx k^* \), where \( k \) is the network’s average degree and \( k^* \) is the wanted average degree. Finally, we add edges from \( i \) permutations of the network’s nodes to the unpermutated nodes, with \( i \) being the smallest integer with \( i > h \), to ensure that every node has a chance of being activated and participating in the network’s dynamics.

Our algorithm has two parameters, \( \delta \) and \( m \), that enable us to tune the degree distribution as well as the clustering coefficient. Generally, lower \( \delta \) and higher \( m \) lead to higher clustering coefficients and scale-free degree distributions, whereas higher \( \delta \) and lower \( m \) lead to low clustering coefficients and exponentially truncated power law degree distributions, see FIG. 1.
[1] A. Barabási and R. Albert, Science 286, 509 (1999).
[2] Y. Lo, C. Li, and S. Lin, Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom) (2012).