

Supplementary Material: Edge of chaos and avalanches in neural networks with heavy-tailed synaptic weight distribution

Łukasz Kuśmierz,^{1,*} Shun Ogawa,¹ and Taro Toyoizumi^{1,2}

¹Laboratory for Neural Computation and Adaptation, RIKEN Center for Brain Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

²Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan

CONTENTS

I. Binary Gaussian network	1
A. Fully connected network	1
B. Sparse network	2
II. Mean field equations correctly predict the mean activity in the thermodynamic limit	3
III. Active phase is chaotic in random networks of binary neurons	3
IV. Details of the simulations	5
V. Avalanche statistics	5
A. Avalanche size distribution	5
B. Avalanche lifetime distribution	6
VI. Existence of a continuous transition in Cauchy networks with a positive threshold	6
References	7

I. BINARY GAUSSIAN NETWORK

A. Fully connected network

For a general activation function it is convenient to describe the behavior of the Gaussian networks in terms of another order parameter,

$$q_0(t) = \frac{1}{N} \sum_i \phi(x_i(t))^2. \quad (\text{S.1})$$

However, in the binary case analyzed here it is equivalent to $m(t)$. The corresponding dynamical mean-field equation [1] reads

$$m(t+1) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\theta}{\sqrt{2m(t)g}} \right) \right], \quad (\text{S.2})$$

where the error function is given by $\operatorname{erf}(x) = 2\pi^{-1/2} \int_0^x \exp(-z^2) dz$. Expanding the RHS of (S.2) around $m(t) \approx 0$ gives

$$m(t+1) = \sqrt{\frac{m(t)g^2}{2\pi\theta^2}} \exp\left(-\frac{\theta^2}{2m(t)g^2}\right) \left(1 + O\left(\frac{m(t)g^2}{\theta^2}\right)\right). \quad (\text{S.3})$$

Due to the exponential factor $m(t+1) < m(t)$ for small enough $m(t)$, which proves that the fixed point $m(t) = 0$ is locally stable under the evolution (S.2). Since the quiescent state is always stable, the transition to chaos, if present, must be discontinuous. This is confirmed by the graphical inspection of (S.2) (Fig. S.1) and in computer simulations (Fig. S.2).

* nalewkoz@gmail.com

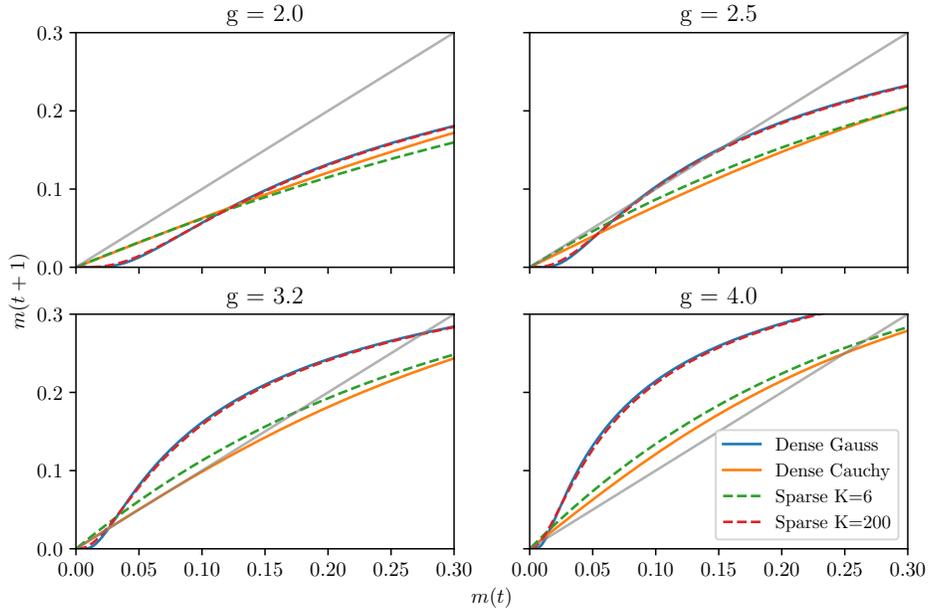


FIG. S.1. Mean field $m(t) \mapsto m(t+1)$ mapping for fully connected (dense) Gaussian and Cauchy networks and sparse Gaussian network with a fixed number of incoming connections per neuron K . For large values of K , the MF equation looks similar to the fully connected Gaussian case. In particular, we observe a discontinuous transition at $g \approx 2.5$. (note that this similarity only holds if g is not too large, since for any finite K there is a continuous transition at which the trivial fixed point loses its stability). For an intermediate sparsity ($2 < K \lesssim 12$), only a second order transition is observed and the dynamics looks qualitatively similar to the dense Cauchy case. For $K \leq 2$ no transition to chaos is observed and the trivial fixed point is always stable (results not shown).

The mapping to a branching process offers a simple way of understanding this result. The probability that a given synapse is an autocrat reads

$$\text{Prob}(J_{ij} > \theta) = \frac{1}{2} \left(1 - \text{erf} \left(-\sqrt{\frac{N}{2}} \frac{\theta}{g} \right) \right) \quad (\text{S.4})$$

In the thermodynamic limit the average number of autocrat connections per neuron can be calculated as before as $\lambda = \lim_{N \rightarrow \infty} NP(J_{ij} > \theta)$, which in our case leads to

$$\lambda = \lim_{N \rightarrow \infty} \sqrt{\frac{N}{2}} \frac{g}{\theta} \exp \left(-\frac{N\theta^2}{2g^2} \right) (1 + O(1/N^2)) = 0, \quad (\text{S.5})$$

i.e. an activity starting from a single seed almost surely dies out.

Interestingly, the discontinuous character of the transition between quiescent and active phases have already been reported in various models that employ Gaussian weights [2, 3], which confirms the generic nature of this result.

B. Sparse network

Let each neuron receive exactly K incoming connections, randomly chosen from the network, and let $J_{ij} \sim \mathcal{N}(0, g^2/K)$. If the mean activity of the network at time t is $m(t)$, the probability that exactly n incoming neurons are active is given by the binomial distribution

$$P(n; m(t)) = \binom{K}{n} m(t)^n (1 - m(t))^{K-n}. \quad (\text{S.6})$$

The membrane potential of a neuron x , conditioned on n incoming connections being active, is a normal random variable with $\mu = 0$ and $\sigma^2 = ng^2/K$. The mean activity of the network in the next step is equal to the probability that the membrane potential of any given neuron crosses the threshold, and so it reads

$$m(t+1) = 2^{-1} \sum_{n=1}^K \binom{K}{n} m(t)^n (1 - m(t))^{K-n} \left(1 - \text{erf} \left(\frac{\theta \sqrt{K}}{\sqrt{2ng}} \right) \right). \quad (\text{S.7})$$

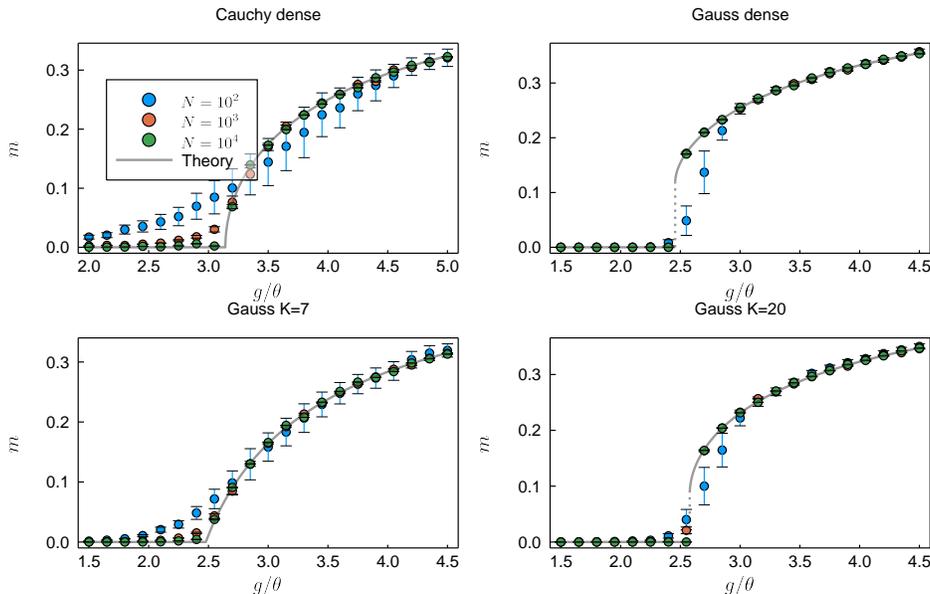


FIG. S.2. Steady-state mean activity as a function of the control parameter g/θ . Lines were obtained by solving our mean field equations (8), (S.2), and (S.7) self-consistently. The points were obtained from computer simulations of $M = 10^5/N$ independent realizations of \mathbf{J} . For each realization of the weight matrix, the network was first evolved for 400 steps from an active state ($m(0) \approx 0.5$). The following 200 steps were used in calculating the average, which was performed over the steps, neurons, and realizations of \mathbf{J} . Error bars denote $\pm 3\sigma$ confidence intervals.

Our computer simulations corroborate the validity of (S.7), see Fig. S.2. Note, additionally, that it is not difficult to show that (S.7) simplifies to the dense case (S.2) in the limit of $K \rightarrow \infty$. Analogous calculations in the case of stochastic units were presented in [4], where the range of validity of the annealed approximation is also discussed.

II. MEAN FIELD EQUATIONS CORRECTLY PREDICT THE MEAN ACTIVITY IN THE THERMODYNAMIC LIMIT

We have performed computer simulations of Cauchy, fully connected Gaussian, and sparse Gaussian networks with binary activation functions (Fig. S.2). The good match between our theoretical predictions and the simulation results strongly suggests that the mean field approach is exact in the thermodynamic limit.

III. ACTIVE PHASE IS CHAOTIC IN RANDOM NETWORKS OF BINARY NEURONS

Here we show, with a simple mathematical argument and computer simulations, that whenever the mean network activity m is non-zero in the steady-state, the corresponding attractor is chaotic. This is true for both Gaussian and Cauchy networks with the binary activation function, and is expected to hold in more general settings, unless some non-trivial structure of the connectivity is present, e.g. symmetric or antisymmetric connectivity (the latter with a symmetric version of the binary activation function, i.e. with $\theta = 0$ and taking values ± 1) leads to non-chaotic, periodic attractors [5].

It is easy to prove our result in the Gaussian case. In this case, with $N \rightarrow \infty$, the maximal Lyapunov exponent is given by [1, 6, 7]

$$\lambda_{\max}^{\text{gauss}} = \frac{1}{2} \ln g^2 \int_{-\infty}^{\infty} Dx [\phi'(g\sqrt{q_0}x)]^2 \quad (\text{S.8})$$

where Dx denotes the standard Gaussian measure, q_0 is the average squared activity (S.1). It is clear that whenever ϕ features any discontinuity and the phase is active ($g\sqrt{q_0} \neq 0$), formula (S.8) predicts $\lambda_{\max}^{\text{gauss}} = \infty$. In particular, the binary activation function has a discontinuity at $x = \theta$ and thus the corresponding active phase is always chaotic. The analysis for the binary Cauchy network is more involved and will be published elsewhere. However, the intuition is the same: small perturbations in the active phase are expanded due to the discontinuity of the activation function. Note that these results are in line with previous

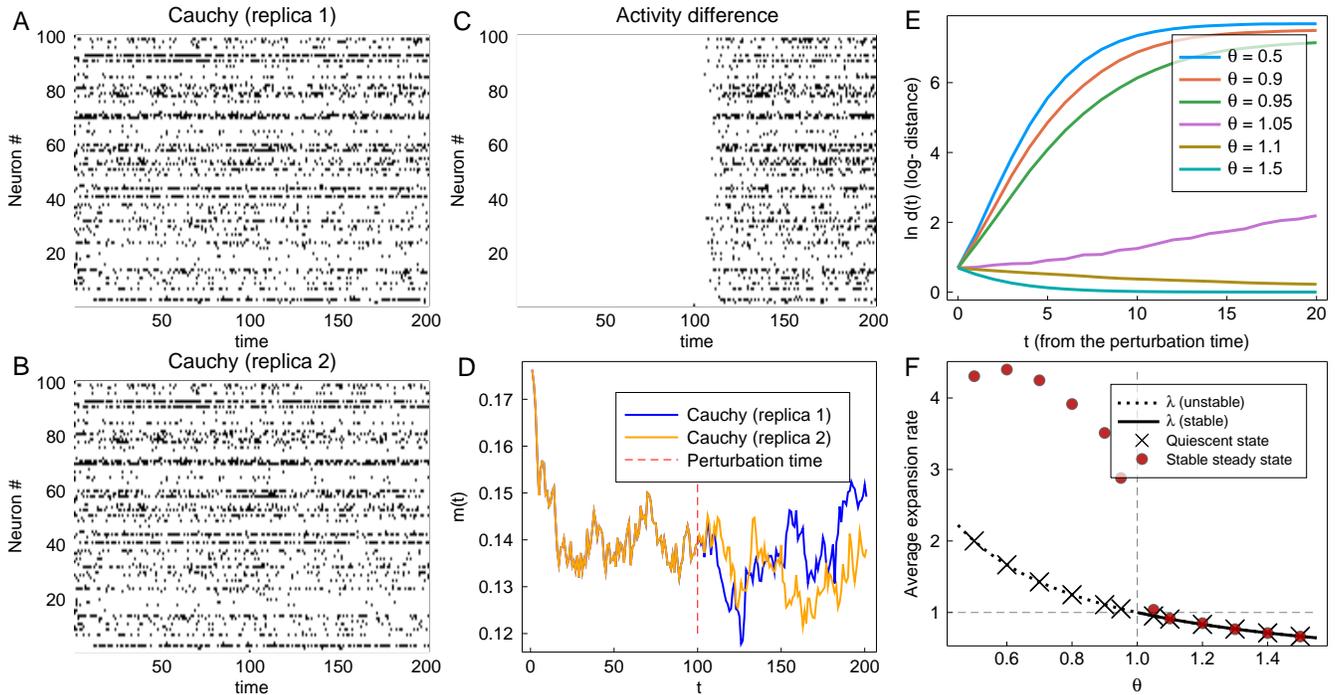


FIG. S.3. Numerical perturbation analysis of the binary Cauchy network. (A-D): Two replicas were evolved from identical initial conditions ($\theta = 0.95$). A single-neuron perturbation, introduced at $T_0 = 100$, is quickly expanded, which can be seen by comparing individual neuron activities (C) or mean network activities (D). (E): The average Hamming distances between perturbed and unperturbed states as functions of the number of steps T after introducing a single-neuron state flip to a steady-state trajectory. The network is sensitive to the perturbations for $\theta < 1$, corroborating the existence of a transition to chaos at the predicted value of $\theta = 1$. (F): The average expansion rate as a function of θ . As expected, the average expansion rate from the quiescent initial state is equal to $\lambda = g/(\pi\theta)$. The average expansion rate from the steady state is different from λ for $\theta < 1$; in this regime the quiescent state does not correspond to the stable steady state (the quiescent state is unstable). In (E) and (F) the distances are averaged over all possible neurons and over 10 realizations of \mathbf{J} . See text for more details.

studies, e.g., it has been shown that in random Boolean networks Lyapunov exponents scale as $\ln K$ with the number of incoming connections [8]. The issue of infinite Lyapunov exponents in binary networks is also discussed in detail in [9].

To support our claim we simulated Cauchy networks with binary activation functions. Two replicas that shared \mathbf{J} were evolved from identical initial conditions. At a fixed time $T_0 = 100$ a perturbation was introduced by flipping the state of a single neuron in one of the replicas. We then calculate the activity difference

$$\delta_i(t) = \left| \phi(x_i^1(t)) - \phi(x_i^2(t)) \right|, \quad (\text{S.9})$$

the Hamming distance

$$d(t) = \sum_{i=1}^N \delta_i(t), \quad (\text{S.10})$$

and the average expansion rate

$$r = \frac{d(T_0 + 1)}{d(T_0)}, \quad (\text{S.11})$$

where $d(T_0) = 1$. Similarly, we calculated the average expansion rate from the quiescent initial state λ . In this case we evolved the (unperturbed) network from a quiescent initial state, which corresponds to the stable steady state for $\theta > \theta^* = 1$ ($g = \pi$), whereas for $\theta < 1$ is unstable. The theoretically calculated branching parameter agrees with λ obtained from the simulations (Fig. S.3F). The evolution of the network is sensitive to the perturbation for $\theta < 1$, which can be observed at the level of individual neuron's activities (Fig. S.3A-C) and, due to the finite size of the network, the mean activity (Fig. S.3D). In the chaotic phase, the Hamming distance initially grows exponentially and then saturates at $d(\infty) \sim N$. The chaotic phase is signaled by $r > 1$, which is directly related to the maximal Lyapunov exponent being positive [8–10]. Although r in general is not equal to λ , both quantities cross the value 1 at the critical point $\theta = 1$. The discrepancy observed around the transition point (the slowly growing distance at $\theta = 1.05$ in Fig. S.3E and the corresponding $r \approx 1.04 > 1$ (Fig. S.3F)) is expected to be a finite size effect.

IV. DETAILS OF THE SIMULATIONS

Computer simulations of binary networks were performed using custom-written codes in MATLAB and Julia [11]. The results for spiking neural networks were obtained with NEST Simulator [12] through a custom-written code in Python. Details of the simulations:

Fig. 2: Networks had size $N = 10^4$ and results were averaged over 10 independent realizations of \mathbf{J} . For a given realization of \mathbf{J} , we run 10^4 simulations, each starting from different seed neuron (i.e. one neuron active, all other neurons inactive). $g = \pi$.

Fig. 3: Both Gaussian and Cauchy networks were fully connected with $N = 10^4$. The injected current changed between -400 pA and 400 pA in small increments every 5 ms. All neurons were of type *iaf_psc_alpha*, which denotes a leaky integrate-and-fire neuron with alpha-shaped postsynaptic currents. Default parameters of the model neuron were used, i.e.: resting potential $E_L = -70$ mV, capacity of the membrane $C_m = 250$ pF, membrane time constant $\tau_m = 10$ ms, refractory period $t_{ref} = 2$ ms, spike threshold $V_{th} = -55$ mV, reset potential $V_{reset} = -70$ mV, rise time of the excitatory and inhibitory synaptic alpha function $\tau_{syn} = 2$ ms. Static synapses were used with the default delay of 1 ms, and weights were randomly drawn from symmetric (a) Gaussian distribution with $\sigma = 2.4 \times 10^3 / \sqrt{N}$ pA, and (b) Cauchy distribution with $\gamma = 1.92 \times 10^3 / N$ pA. Poisson noise was injected randomly into the network, activating each neuron approximately twice every second. Note that without any external input the analyzed networks cannot recover from the quiescent state, since the model neurons are never spontaneously active. Sub-sampling: For the sake of clarity and drawing efficiency, in the raster plots only 10% of spikes of 100 randomly chosen neurons were drawn. Activity histograms were created using all data.

Fig. S.3: The averaging was performed over all $N = 10^4$ neurons (i.e., including the unperturbed network, $N + 1$ replicas were simulated) and over 10 realizations of \mathbf{J} . At time $T = 0$ the unperturbed network was prepared in the (a) steady state by evolving it for 100 steps before introducing perturbations or (b) quiescent state. $g = \pi$. Sub-sampling: For the sake of clarity and drawing efficiency, activities of a randomly chosen subpopulation of neurons ($N' = 100$) is shown.

V. AVALANCHE STATISTICS

The mapping to the branching process together with the general results known for critical branching processes provide the critical exponents [13–15]. Here, for completeness, we calculate two critical exponents in our specific case.

A. Avalanche size distribution

The size of an avalanche is defined as the sum of the number of active neurons at each time step, from the beginning of the avalanche till its end. Let S_m denote the size of an avalanche starting from m seeds and $G_m(z)$ denote the corresponding generating function

$$G_m(z) \equiv \langle z^{S_m} \rangle = \sum_{j=0}^{\infty} z^j \text{Prob}(S_m = j). \quad (\text{S.12})$$

Since the activity of network is assumed to be sparse, the avalanche that starts from m seeds consists of m independent avalanches starting from a single seed. Therefore we can write that $S_m = \sum_{i=1}^m S_1^{(i)}$, where $\{S_1^{(i)}\}$ is a set of i.i.d. random variables denoting sizes of single-seed generated avalanches. At the level of the generating functions, this assumption leads to a simple expression

$$G_m(z) = [G_1(z)]^m. \quad (\text{S.13})$$

On the other hand, we know how a seed neuron propagates the activity through the network in a single step: it activates m neurons with probability

$$p_m = \frac{\lambda^m}{m!} e^{-\lambda}, \quad (\text{S.14})$$

where λ is given by (10). This means that one seed generate an avalanche of size $1 + S_m$ with probability p_m , where 1 is from the first step (i.e., the seed) and S_m is from the subsequent steps. Hence, the single-seed generating function can be calculated as follows

$$G_1(z) = \sum_{m=0}^{\infty} p_m \langle z^{1+S_m} \rangle = z \sum_{m=0}^{\infty} p_m G_m(z). \quad (\text{S.15})$$

We combine (S.15) with (S.13) and (S.14) and arrive at an implicit expression for the one-seed avalanche size generating function

$$G_1(z) = z \left\{ p_0 e^{S_0} + \sum_{m=1}^{\infty} p_m [G_1(z)]^m \right\} = z \exp(\lambda G_1(z) - \lambda). \quad (\text{S.16})$$

Note that we used $S_0 = 0$ (no avalanche without a seed). In order to inspect the tail of the distribution of S_1 we introduce an auxiliary function $g(\epsilon) = 1 - G_1(1 - \epsilon)$ and expand the RHS of (S.16) assuming that $\epsilon \ll 1$ and $g(\epsilon) \ll 1$ (valid for $\lambda \leq 1$):

$$1 - g(\epsilon) = (1 - \epsilon) \left(1 - \lambda g(\epsilon) + \frac{\lambda^2 g(\epsilon)^2}{2} + O(g(\epsilon)^3) \right), \quad (\text{S.17})$$

which to the lowest order can be rewritten as

$$g(\epsilon) = \begin{cases} \frac{\epsilon}{1-\lambda}, & \text{for } \lambda \neq 1 \\ \sqrt{2\epsilon}, & \text{for } \lambda = 1. \end{cases} \quad (\text{S.18})$$

The small ϵ behavior of $g(\epsilon) \sim \epsilon^{1/2}$ in the $\lambda = 1$ case translates into the tail behavior of the avalanche size density as

$$\text{Prob}(S_1 = s) \sim s^{-3/2} \quad (\text{S.19})$$

for large s .

B. Avalanche lifetime distribution

Let T_m be the lifetime (number of steps with nonzero activity) of an avalanche that starts from m seeds. By definition $T_0 = 0$ and $T_m \geq 1$ for $m > 0$. As before, we treat an avalanche from different seeds as independent, and thus the following identity linking the survival probabilities holds

$$Q_m(t) = 1 - \text{Prob}(T_m \leq t) = 1 - [\text{Prob}(T_1 \leq t)]^m = 1 - [1 - Q_1(t)]^m. \quad (\text{S.20})$$

As in the case of the size distribution, we can unwrap the first step of the dynamics starting from a single seed, which gives

$$Q_1(t+1) = \sum_{m=1}^{\infty} p_m Q_m(t) = 1 - \sum_{m=0}^{\infty} p_m [1 - Q_m(t)]. \quad (\text{S.21})$$

We plug (S.20) and (S.14) into (S.21) and arrive at the following recursive relation

$$Q_1(t+1) = 1 - \exp[-\lambda Q_1(t)]. \quad (\text{S.22})$$

If $\lambda > 1$ there exists a non-zero fixed point corresponding to the non-zero probability of survival at $t \rightarrow \infty$. In contrast, the activity eventually dies out almost surely for $\lambda \leq 1$. Assuming $\lambda Q_1(t) \ll 1$ the recursive relation simplifies to

$$Q_1(t+1) = \lambda Q_1(t) - \frac{\lambda^2}{2} Q_1(t)^2, \quad (\text{S.23})$$

which predicts an exponential decay for $\lambda < 1$. At the critical point $\lambda = 1$ and the decay is a power law. In that case the recursion can be solved with an ansatz $Q_1(t) = C/t^\delta$, leading to $\delta = 1$, as expected.

VI. EXISTENCE OF A CONTINUOUS TRANSITION IN CAUCHY NETWORKS WITH A POSITIVE THRESHOLD

Let $\phi(x)$ be an activation function such that $\phi(x) = 0$ for x below a positive threshold θ , and $\phi(x) \approx C$ for sufficiently large x (i.e. $x > m_1$). Without much loss of generality, we additionally assume that $\phi(x) \geq 0$ and $\int_a^b \phi(x) dx < \infty$. The integral in the mean-field equation can be then decomposed as follows:

$$m(t+1) = \frac{1}{\pi} \int_0^{\infty} dz \frac{\phi(m(t)gz)}{1+z^2} \approx \frac{1}{\pi} \int_{\frac{\theta}{m(t)g}}^{\frac{m_1}{m(t)g}} dz \frac{\phi(m(t)gz)}{1+z^2} + \frac{C}{\pi} \int_{\frac{m_1}{m(t)g}}^{\infty} dz \frac{1}{1+z^2} = \frac{m(t)g}{\pi} \int_{\theta}^{m_1} dy \frac{\phi(y)}{[m(t)g]^2 + y^2} + \frac{C}{\pi} \arctan\left(\frac{m(t)g}{m_1}\right). \quad (\text{S.24})$$

We expand (S.24) around $m(t) = 0$:

$$m(t+1) = \frac{m(t)g}{\pi} \left(\int_{\theta}^{m_1} dy \frac{\phi(y)}{y^2} + \frac{C}{m_1} \right) + O(m(t)^3), \quad (\text{S.25})$$

and conclude that the transition between quiescent and active state occurs at the critical point described by the equation

$$\frac{g}{\pi} \left(\int_{\theta}^{m_1} dy \frac{\phi(y)}{y^2} + \frac{C}{m_1} \right) = 1. \quad (\text{S.26})$$

Hence, this guarantees the existence of positive and finite critical g^* that solves the above equation. It is possible to extend these results to activation functions $\phi(x)$ that are non-zero around $x = 0$ and are non-saturating: the transition exists if $\phi(x)$ is sufficiently superlinear around $x = 0$ and sublinear for large $|x|$.

-
- [1] L. Molgedey, J. Schuchhardt, and H. G. Schuster, *Physical review letters* **69**, 3717 (1992).
 - [2] N. Brunel, *Journal of computational neuroscience* **8**, 183 (2000).
 - [3] J. H. Marshel, Y. S. Kim, T. A. Machado, S. Quirin, B. Benson, J. Kadmon, C. Raja, A. Chibukhchyan, C. Ramakrishnan, M. Inoue, *et al.*, *Science* **365**, eaaw5202 (2019).
 - [4] B. Derrida, E. Gardner, and A. Zippelius, *EPL (Europhysics Letters)* **4**, 167 (1987).
 - [5] E. Goles, *Discrete Applied Mathematics* **13**, 97 (1986).
 - [6] M. Massar and S. Massar, *Physical Review E* **87**, 042809 (2013).
 - [7] G. Wainrib and M. N. Galtier, *Neural Networks* **76**, 39 (2016).
 - [8] B. Luque and R. V. Solé, *Physica A: Statistical Mechanics and its Applications* **284**, 33 (2000).
 - [9] C. v. Vreeswijk and H. Sompolinsky, *Neural computation* **10**, 1321 (1998).
 - [10] L. Büsing, B. Schrauwen, and R. Legenstein, *Neural computation* **22**, 1272 (2010).
 - [11] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *SIAM review* **59**, 65 (2017).
 - [12] C. Linssen, M. E. Lepperød, J. Mitchell, J. Pronold, J. M. Eppler, C. Keup, A. Peyser, S. Kunkel, P. Weidel, Y. Nodem, D. Terhorst, R. Deepu, M. Deger, J. Hahne, A. Sinha, A. Antonietti, M. Schmidt, L. Paz, J. Garrido, T. Ippen, L. Riquelme, A. Serenko, T. Kühn, I. Kitayama, H. Mørk, S. Spreizer, J. Jordan, J. Krishnan, M. Senden, E. Hagen, A. Shusharin, S. B. Vennemo, D. Rodarie, A. Morrison, S. Graber, J. Schuecker, S. Diaz, B. Zajzon, and H. E. Plesser, “Nest 2.16.0,” (2018).
 - [13] P. Alstrøm, *Physical Review A* **38**, 4905 (1988).
 - [14] M. A. Muñoz, R. Dickman, A. Vespignani, and S. Zapperi, *Physical Review E* **59**, 6175 (1999).
 - [15] G. Ódor, *Reviews of modern physics* **76**, 663 (2004).