BOOTSTRAP PERCOLATION ON A RANDOM GRAPH
COUPLED WITH A LATTICE

SVANTE JANSON, ROBERT KOZMA, MIKLÓS RUSZINKÓ, AND YURY SOKOLOV

ABSTRACT. In this paper a random graph model $G_{\mathbb{Z}_N^2, p_d}$ is introduced, which is a combination of fixed torus grid edges in $(\mathbb{Z}/N\mathbb{Z})^2$ and some additional random ones. The random edges are called long, and the probability of having a long edge between vertices $u, v \in (\mathbb{Z}/N\mathbb{Z})^2$ with graph distance d on the torus grid is $p_d = c/Nd$, where c is some constant. We show that, *whp*, the diameter $D(G_{\mathbb{Z}_N^2, p_d}) = \Theta(\log N)$. Moreover, we consider non-monotonous bootstrap percolation on $G_{\mathbb{Z}_N^2, p_d}$. We prove the presence of phase transitions in mean-field approximation and provide fairly sharp bounds on the error of the critical parameters. Our model addresses interesting mathematical questions of non-monotonous bootstrap percolation, and it is motivated by recent results of brain research.

1. INTRODUCTION

Bootstrap percolation on lattices has been extensively investigated in the last decades, and recently it has been considered on the classical Erdős-Rényi random graph $G_{n,p}$ as well [14]. In this paper we consider a stochastic process of activation propagation over the random graph which combines lattice $\mathbb{Z}^2$ with additional random edges that depend on the distance between vertices. A similar graph has been studied, by, e.g., Aizenman, Kesten and Newman [1] - the so-called long-range percolation graph. In that model a pair of sites of d-dimensional lattice $\mathbb{Z}^d$ is connected (or a bond is occupied) with probability that depends on the graph distance. In this paper we change the way probabilities are defined to get a sparser graph with respect to long edges.

The n-cycle long-range percolation graph (a cycle lattice with n vertices) was considered in [4] where the probabilities of random edges decay polynomially as their distances increase. It was shown that the diameter of this graph is of the order of $\log N$, assuming that the parameters are constrained to a certain parameter region. The combination of a d-dimensional grid with certain random edges (decreasing in distance) was considered by Coppersmith, Gamarnik and Sviridenko [9]. They showed that under certain conditions on the dimension and probability p, the diameter is either $\Theta(\log N)$ or N^η, where the power coefficient η satisfies $0 < \eta < 1$. Bollobás and Chung [6] investigated the combination of an n-cycle with a random matching. Clearly, the n-cycle has diameter $\lfloor n/2 \rfloor$. However, a random

---

matching reduces the diameter drastically and it becomes $(1 + o(1)) \log_2 n$. Later, Watts and Strogatz [21] introduced the "*small world*" model on the vertex set of the n-cycle, where the edges are *rewired* at random with probability p, starting from a circle lattice with n vertices and k edges per vertex. This construction allows to 'tune' the graph between regularity ($p = 0$) and disorder ($p = 1$). A different version of the "*small world*" model has been described by Newman and Watts [18]. Here, again an n-cycle was considered and the edges of the cycle were fixed. In contrast to the original formulation [21], however, in [18] random edges were added with some probability instead of rewiring the edges of the cycle. This approach results in a smaller diameter for a given graph size. There has been a lot of interest in studying scaling behavior and phase transitions in inhomogeneous random graph models, see, for example [7].

In [14], a theory of bootstrap percolation has been developed on $G_{n,p}$ regarding the size a of the set of initially active sites. Results include sharp threshold for phase transition for parameters p and a, and results for the time t required to the termination of the bootstrap percolation process.

Recently, Turova and Vallier [20] considered bootstrap percolation on the combination of the lattice $\mathbb{Z}^d$ and the random graph $G_{n,p}$, where the edges of $\mathbb{Z}^d$ and $G_{n,p}$ are selected with probability q and p, respectively. Sharp threshold for phase transition was derived. The authors got asymptotic results for the time when the bootstrap percolation process stops.

Here we consider a random graph G that is built as follows. We start with the $\mathbb{Z}^2$ lattice over a $(N + 1) \times (N + 1)$ grid; for the sake of simplicity we assume periodic boundary conditions. Thus, we have a torus $\mathbb{T}^2 = (\mathbb{Z}/N\mathbb{Z})^2$, with the short notation $\mathbb{Z}_N^2$. The set of vertices of G consists of all vertices of $\mathbb{Z}_N^2$, in total N^2 vertices. All the edges from the torus $\mathbb{Z}_N^2$ are included in the graph G. In addition, we introduce random edges as follows. For every pair of vertices we assign an edge with probability that depends on the graph distance d between the two vertices, i.e., d is the length of the shortest path between the given pair of vertices in the torus grid. Accordingly, the probability of a long edge is described as follows:

$$(1.1) \qquad p_d = \mathbb{P}\left((u, v) \in E(G) \mid dist(u, v) = d\right) = \frac{c}{N} \times d^{-\alpha},$$

where c and α are positive constants, $d > 1$ (no multiple edges are allowed between any pair of vertices) and N is large enough so that each $p_d < 1$. We assume $\alpha = 1$ throughout this study. We will denote this model the $G_{\mathbb{Z}_N^2, p_d}$ graph. The edges of the torus are called *short edges*, while the randomly added ones are called *long edges*.

The model introduced in this paper is also motivated by the structure and operation of the neuropil, the densely connected neural tissue of the cortex [11, 15]. The human brain has about 10^{11} neurons. Typically, a neuron has several thousands of connections to other neurons through synapses, thus the human brain has $\sim 10^{15}$ synaptic connections. Most of the connections are short and limited to the neuron's direct neighborhood (in some metric), forming the so-called the *dendritic arbor*. In addition, the neurons have a few long connections (axons), which extend further away from their cell body. In general, there are several thousands short connections in the dendritic arbor for one distant connection represented by a long axon. We use $G_{\mathbb{Z}_N^2, p_d}$ to model the combined effect of mostly short connections and a few long connections. It is more likely to have in brains shorter connections than

longer ones, which is a fact captured in the definition of p_d, as p_d is decreasing in the graph distance d.

There are two types of neurons in the brain, namely excitatory and inhibitory ones. The type of a neuron describes the function of the neuron in the brain, e.g., excitatory (inhibitory) neurons excite (inhibit) the neurons to which they are connected. It is known that there are much more excitatory neurons than inhibitory neurons in the cortex; the ratio of inhibitory to excitatory neurons is typically $1/4$ [12]. Based on neuroscience studies it is expected that pure excitatory populations can maintain non-zero background activation level, while interacting excitatory and inhibitory populations are able to produce limit cycle oscillations [11, 15]. This paper focuses on conditions required to sustain non-zero activity level in pure excitatory network.

The present work is organized as follows: first we describe some properties of the introduced random graph $G_{\mathbb{Z}_N^2, p_d}$. We derive bounds on the diameter of this graph and describe its degree distribution. The second part of this paper is devoted to the study of activation processes. To simplify the mathematical treatment we analyze the activation as a stochastic process in mean-field approximation [2], for which we derive conditions for phase transitions in the presence of single type of (excitatory) nodes. Activity propagation in graphs with two types of nodes (excitatory and inhibitory) is the objective of future studies [16].

We will use the following standard notation; for non-negative sequences a_m and b_m, $a_m = O(b_m)$ if $a_m \leq c b_m$ holds for some constant $c > 0$ and every m; $a_m = \Theta(b_m)$ if both $a_m = O(b_m)$ and $b_m = O(a_m)$ hold; $a_m \sim b_m$ if $\lim_{m \to \infty} a_m / b_m = 1$; $a_m = o(b_m)$ if $\lim_{m \to \infty} a_m / b_m = 0$. A sequence of events A_n occurs with high probability, *whp*, if the probability $\mathbb{P}(A_n) = 1 - o(1)$.

2. Properties of $G_{\mathbb{Z}_N^2, p_d}$

First notice that the expected number of long edges $E_\ell \subseteq E(G_{\mathbb{Z}_N^2, p_d})$ is proportional to N^2.

Claim 1. $\mathbb{E}(|E_\ell|) \sim (2c \ln 2)N^2, \quad i.e., \quad \lim_{N \to \infty} \dfrac{\mathbb{E}(|E_\ell|)}{2cN^2 \ln 2} = 1.$

Proof. Indeed, the number of vertices $|\Lambda_d|$ in $\mathbb{Z}_N^2$ which are exactly at distance d from a fixed vertex is

$$|\Lambda_d| = \begin{cases} 4d, & 1 \leq d \leq \lfloor N/2 \rfloor \\ 4(N - d), & \lfloor N/2 \rfloor < d \leq N \end{cases}$$

for N odd, and

$$|\Lambda_d| = \begin{cases} 4d, & 1 \leq d < N/2 \\ 4d - 2, & d = N/2 \\ 4(N - d), & N/2 < d < N \\ 1, & d = N \end{cases}$$

for N even. The number of pairs of vertices in $\mathbb{Z}_N^2$ having distance d is $\frac{N^2 |\Lambda_d|}{2}$. Therefore, for N odd

$$\mathbb{E}(|E_\ell|) = \sum_{d=2}^{N} \frac{N^2 |\Lambda_d|}{2} \frac{c}{Nd} = \sum_{d=2}^{N/2} \frac{4N^2 d}{2} \frac{c}{Nd} + \sum_{d=N/2+1}^{N} \frac{4N^2(N - d)}{2} \frac{c}{Nd}$$

$$(2.1) \qquad = \quad (2c \ln 2)N^2 + O(N) \sim (2c \ln 2)N^2.$$

For N even a similar computation gives the same result. $\qquad \square$

2.1. **Degree distribution.** The degree distribution of a vertex $v \in G_{\mathbb{Z}_N^2, p_d}$ with respect to long edges can be approximated by Poisson distribution. Let W be the random variable describing the degree of a particular vertex v considering long edges only. Then clearly, the degree of a vertex $v \in G_{\mathbb{Z}_N^2, p_d}$ considering the short edges, too, is $W + 4$.

Lemma 1. *The probability that a vertex has degree k considering only the long edges is given by*

$$(2.2) \qquad \mathbb{P}(W = k) = \sum_{k_2 + \ldots + k_N = k} \prod_{i=2}^{N} \binom{|\Lambda_i|}{k_i} \left(\frac{c}{Ni}\right)^{k_i} \left(1 - \frac{c}{Ni}\right)^{|\Lambda_i| - k_i}.$$

The total variation distance

$$(2.3) \qquad d_{TV}\left(\mathcal{L}(W), Po(\lambda)\right) = \frac{1}{2} \sum_{j \geq 0} |\mathbb{P}(W = j) - \mathbb{P}(Y = j)| = O(1/N),$$

where the random variable Y has Poisson distribution $Po(\lambda)$, with $\lambda = 4c \ln 2$.

Proof. The probability of the event A_i that a vertex has k_i edges of length i is clearly

$$(2.4) \qquad \mathbb{P}(A_i) = \binom{|\Lambda_i|}{k_i} \left(\frac{c}{Ni}\right)^{k_i} \left(1 - \frac{c}{Ni}\right)^{|\Lambda_i| - k_i}$$

Therefore, the probability that a vertex has degree exactly k is

$$(2.5) \quad \mathbb{P}(W = k) = \mathbb{P}\left(\bigcup_{k_2 + \ldots + k_N = k} \bigcap_{i=2}^{N} A_i\right) = \sum_{k_2 + \ldots + k_N = k} \prod_{i=2}^{N} \mathbb{P}(A_i)$$

$$= \sum_{k_2 + \ldots + k_N = k} \prod_{i=2}^{N} \binom{|\Lambda_i|}{k_i} \left(\frac{c}{Ni}\right)^{k_i} \left(1 - \frac{c}{Ni}\right)^{|\Lambda_i| - k_i}.$$

The last expression is not very convenient to use. However, a standard Poisson approximation can be given using Le Cam's argument [17], see also e.g. [3]. Pick an arbitrary vertex v and let enumerate the other $N^2 - 5$ vertices by u_i, $i = 1, \ldots, N^2 - 5$, excluding the nearest neighbors, i.e., vertices at distance 1. The long edges that connect the vertex v to other vertices of the graph are independent 0–1 random variables with Bernoulli $Be(p_i)$ distribution. In other words, let $I_i = 1$ be the event that there is an edge between vertices v and u_i, so that $\mathbb{P}(I_i = 1) = p_i$ and $\mathbb{P}(I_i = 0) = 1 - p_i$, where p_i may in general vary for different i. Consider now the degree $W = \sum_{i=1}^{N^2 - 5} I_i$ of the vertex v. Let

$$\lambda_1 = \sum_{i=1}^{N^2 - 5} p_i = 4c \ln 2 + O(1/N),$$

where the last equality follows from Eq. (2.1). By triangle inequality,

$$(2.6) \qquad d_{TV}\left(\mathcal{L}(W), \text{Po}(\lambda)\right) \leq d_{TV}\left(\mathcal{L}(W), \text{Po}(\lambda_1)\right) + d_{TV}\left(\text{Po}(\lambda_1), \text{Po}(\lambda)\right)$$

The first term, by Le Cam [17], see also [3, Theorem 2.M], is at most

$$\sum_{i=1}^{N^2-5} p_i^2 \;=\; \sum_{d=2}^{N} |\Lambda_d| p_d^2 \leq \sum_{d=1}^{N} |\Lambda_d| p_d^2 = \sum_{d=1}^{N/2} 4d \left(\frac{c}{Nd}\right)^2$$

$$(2.7) \qquad + \sum_{d=N/2+1}^{N} 4(N-d) \left(\frac{c}{Nd}\right)^2 \leq \sum_{d=1}^{N} 4d \left(\frac{c}{Nd}\right)^2 = O\left(\frac{\ln N}{N^2}\right).$$

and by Theorem 1.C (i) in [3]

$$(2.8) \qquad\qquad d_{TV}\left(\mathrm{Po}(\lambda_1), \mathrm{Po}(\lambda)\right) = O\left(|\lambda_1 - \lambda|\right) = O\left(\frac{1}{N}\right).$$

$\square$

Clearly, Lemma 1 also implies that in Eq. (2.3) each term satisfies $|\mathbb{P}(W = j) - \mathbb{P}(Y = j)| = O(1/N)$.

2.2. The diameter of $G_{\mathbb{Z}_N^2, p_d}$.
Next we show that the addition of long edges to the torus grid reduces significantly (from linear to logarithmic in the number of vertices) its diameter.

Theorem 1. *There exist constants C_1, C_2, which depend on c only, such that for the diameter $D(G_{\mathbb{Z}_N^2, p_d})$ the following hold.*

$$\lim_{N \to \infty} \mathbb{P}\left(C_1 \log N \leq D(G_{\mathbb{Z}_N^2, p_d}) \leq C_2 \log N)\right) = 1, \; i.e., \; D(G_{\mathbb{Z}_N^2, p_d}) = \Theta(\log N), \; whp.$$

Proof. The lower bound is trivial. The expected degree $\mathbb{E}(d(v))$ of a vertex v by Claim 1 is a constant $k = k(c)$. Thus, the expected number of vertices A_m we can reach in at most $m \geq 0$ steps from a given vertex v is less than or equal to $1 + \sum_{i=1}^{m} k(k-1)^{i-1}$. For $m \geq 3$, this is less than k^m, and thus, by Markov's inequality,

$$(2.9) \qquad\qquad \mathbb{P}(A_m \geq N^2) \leq \frac{\mathbb{E}(A_m)}{N^2} \leq \frac{k^m}{N^2}.$$

If we choose $m \leq C_1 \log N$ with C_1 sufficiently small, the probability in Eq. (2.9) tends to zero, i.e., even from a given vertex v we cannot reach all vertices within distance $C_1 \log N$. Hence, $C_1 \log N$ bounds the diameter from below.

To prove the upper bound, partition the vertices of $G_{\mathbb{Z}_N^2, p_d}$ into consecutive $k \times k$ blocks $B_{ij}, i, j = 1, \ldots, \frac{N}{k}$, where k is a constant $k(c)$ to be chosen later. (For simplicity, we will assume that everywhere divisibility holds during the proof; otherwise we let some blocks be $(k+1) \times (k+1)$.) Define the graph G' as follows. The vertices are the blocks, and two blocks $B_{i,j}$ and $B_{k,\ell}$, $(1, \leq i, j, k, \ell \leq N/k)$ are connected iff there is a long edge from a vertex of $B_{i,j}$ to a vertex of $B_{k,\ell}$ in $G_{\mathbb{Z}_N^2, p_d}$. We obtain a random graph on N^2/k^2 vertices where the edge probabilities can be obtained from the ones of $G_{\mathbb{Z}_N^2, p_d}$. For an arbitrary pair of vertices $B_{i,j}$ and $B_{k,\ell}$, the probability of the event $A_{i,j;k,l}$ that they are connected is bounded from below by the probability, that two blocks which are most distant from each other in $\mathbb{Z}_N^2$ are connected. Therefore, for large N,

$$\mathbb{P}(A_{i,j;k,l}) \;\geq\; \mathbb{P}(A_{1,1;N/(2k),N/(2k)}) = 1 - \mathbb{P}(\overline{A_{1,1,\frac{N}{2k},\frac{N}{2k}}}) \geq 1 - (1 - p_N)^{k^4}$$

$$= \; 1 - \left(1 - \frac{c}{N^2}\right)^{k^4} \geq 1 - e^{-ck^4/N^2} \geq ck^4/2N^2.$$

For the second inequality we picked the two most distant vertices from each block, and the last one follows from $e^x \leq 1 + x/2$ for $x < 0$ sufficiently close to 0.

By, e.g., Theorem 9.b in the seminal paper of Erdős and Rényi [10] there is a constant c_1 such that in the Erdős-Rényi random graph $G_{n,p}$ with $p = c_1/n$ there is a giant component on at least, say, $n/2$ vertices, *whp*. Choosing

$$k \geq (2c_1/c)^{1/2}$$

we get that the probability that G' has an edge for arbitrary pair of vertices is at least

$$ck^4/2N^2 \geq c_1 k^2/N^2.$$

Since the edges of G' are chosen independently, it will contain a giant component on at least $N^2/2k^2$ vertices, *whp*. The diameter of the giant component of $G_{n,p}$ with $p = c_1/n$ is known to be of order $O(\log n)$, *whp*. (See, e.g. Table 1 in [8].)

First, assume that vertices $u, v \in G_{\mathbb{Z}_N^2, p_d}$ are contained in blocks $B(u)$ and $B(v)$ which are vertices of the giant component in G'. Find the shortest path, say, $B(u) = B(x_0), B(x_1), B(x_2), \ldots, B(x_m) = B(v)$, between $B(u)$ and $B(v)$ in G'. Let $(x_0, x_1), (x_1', x_2), (x_2', x_3), \ldots, (x_{m-1}', x_m)$, $x_i, x_i' \in B(x_i)$ be the edges in $G_{\mathbb{Z}_N^2, p_d}$ inducing this path in G'.

Now go from u to x_0 in $B(u)$ along short ($\mathbb{Z}^2$) edges. Jump from x_0 to x_1. Then go from x_1 to x_1' in $B(x_1)$ along short edges. Jump from x_1 to x_2', and so on. The total length of the path from u to v, will be at most

$$m + 2k(m + 1) \leq (2k + 1)(m + 1).$$

Indeed, we make m jumps, and within each block we make at most $2k$ steps along short edges. Since $m = O(\log N)$, *whp*, the proof of this case is finished.

Next we show that, *whp*, every vertex $v \in G_{\mathbb{Z}_N^2, p_d}$ is close to some block B of the giant component in G'. Indeed, by symmetry, the set A of vertices in the giant component can be any set of vertices of the same size, with the same probability. Therefore, one can regard A as a uniformly random subset on at least half of the vertices in G'.

For some large constant D, the number of vertices with distance at most $D\sqrt{\log_2 N}$ from a fixed vertex v in $\mathbb{Z}^2$ is

$$\sum_{d=1}^{D\sqrt{\log_2 N}} 4d \geq 4D^2 \log_2 N,$$

i.e., this neighborhood contains a vertex from at least

$$\frac{4D^2 \log_2 N}{k^2}$$

blocks. Since A contains at least half of the vertices in G', the probability that none of those blocks is in A is

$$\leq 2^{-\frac{4D^2 \log_2 N}{k^2}} = N^{-4D^2/k^2}.$$

Therefore, the probability that there is a vertex $v \in G_{\mathbb{Z}_N^2, p_d}$ for which there is no vertex u within distance $D\sqrt{\log_2 N}$ such that $B(u) \in A$ is

$$\leq N^2 \cdot N^{-4D^2/k^2} < N^{-2},$$

assuming that D is large enough.

Now, consider arbitrary two vertices $u, v \in G_{\mathbb{Z}_N^2, p_d}$. If one, or neither of them is in block from A, then, *whp*, each of them can reach a block from A within $D\sqrt{\log_2 N}$ steps in $\mathbb{Z}^2$, and then proceed as in case $B(u), B(v) \in A$. Since the number of additional steps *whp* is $O(\sqrt{\log N})$, the proof is finished. $\qquad\square$

3. Activation process on the random graph $G_{\mathbb{Z}_N^2, p_d}$

Now we introduce a stochastic process on the graph we have just built. Each vertex is described by two variables, namely, the *type* of the vertex and its *state*. The *type* of the vertex is either excitatory (E) or inhibitory (I). The other variable represents the *state* of the vertex, which can be active or inactive. In other words, for each vertex we attribute a 2-dimensional vector where each coordinate is a Bernoulli random variable. The *type* is selected at the start and it remains unchanged through the process, while the *state* of the vertex changes during the process according to some rules specified next.

Let $A(t)$ denote the set of all active vertices at time t, while $A_E(t)$ and $A_I(t)$ are the sets of active vertices of type E and I at time t, respectively; $A(t) = A_E(t) \cup A_I(t)$. We also define a potential function $\chi_v(t)$ for each vertex v such that $\chi_v(t) = 1$ if vertex v is active at time t, and $\chi_v(t) = 0$ if v is inactive. Therefore, $A_i(t) = \{v \in V(G_{\mathbb{Z}_N^2, p_d}) \mid \chi_v(t) = 1 \ \& \ v \text{ is of type } i\}, i \in \{E, I\}$. At the beginning, let $A(0)$ be a random subset of vertices with each vertex active with probability p, independently of its type and of all other vertices. Each vertex may change its activity based on the states of its neighbors. For a vertex v of type E, the evolution rule is

$$(3.1) \qquad \chi_v(t+1) = \mathbb{1}\left(\sum_{u \in N^E(v)} \chi_u(t) - \sum_{u \in N^I(v)} \chi_u(t) \geq k\right),$$

where $N^E(v)$ and $N^I(v)$ denote the subsets of vertices in the closed neighborhood of the vertex v, of type E and I, respectively; and $\mathbb{1}$ is the indicator function. Here k is a nonnegative integer that specifies a threshold required for the node to be in the active state. For a vertex v of type I, the following rule holds:

$$(3.2)$$
$$\chi_v(t+1) = \mathbb{1}\left(\sum_{u \in N^E(v)} \chi_u(t) + \sum_{u \in N^I(v)} \chi_u(t) \geq k\right) = \mathbb{1}\left(\sum_{u \in N(v)} \chi_u(t) \geq k\right),$$

where $N(v) = N^E(v) \cup N^I(v)$ is the closed neighborhood of vertex v. Notice, that vertices of type E and I have different roles and influence each other differently.

Observe, that the set of active vertices does not necessarily grow monotonically during the activation process even in the absence of nodes of type I, whereas monotonicity is typically assumed in most bootstrap percolation models. In this paper we focus on activation processes having only type E vertices. In this case, according to Eq. (3.1) we have the k-rule, i.e., a vertex will be active at the next time step if it has at least k active neighbors including itself. We assume for simplicity that k is not greater than 3 in the present study. Notice that if there are only local edges, the case $k = 3$ yields *bootstrap percolation* with majority rule. The choice of small k is motivated by the fact that there are vertices with degree 4 with positive probability. Therefore, if k could be 5 or more, then there would be vertices that cannot become active unless they were activated at the beginning.

4. Mean-field approximation for single type of nodes

As we mentioned above, we assume that there is only one type of nodes (E). Moreover, for simplicity, we consider here the *mean-field* (MF) approximation of the model. The mean-field approximation assumes that the activation and degrees of the various nodes are well-mixed; hence we ignore any dependencies between activation and vertex degrees, as well as any dependencies between the state of a vertex and the state of its neighbors [2]. Effectively, we sample a new set of neighbors at each step. This, in particular, implies that the MF approximation does not depend on the topology of the torus but only on the degree distribution and on the cardinality of $A(t)$. Thus, the transition probabilities from one state to another depend only on the number of active nodes. Furthermore, we assume that the vertices are activated independently of each other, ignoring the small dependencies between degrees and activities for different vertices.

4.1. **Phase transition in mean-field model.** Let $\rho_t = A(t)/N^2$, where N^2 is the size of the torus. Clearly, $\rho_t \in [0,1]$ and it defines the density of active nodes at time t.

The mean-field analysis is an analytical approach of finding phase transitions in the stochastic process by averaging the system over space. Thus, the mean-field model reduces the analysis of a system with distributed components to a system with a single component. Let $f(\rho_t)$ denote the conditional mean of ρ_{t+1} given ρ_t, for the mean-field approximation. The main task of the mean-field approximation is to find solutions to the *fixed-point equation* $x = f(x)$, where the solutions of this equation are called fixed points. This approach is based on the observation that the critical behavior of the original system often occurs near the unstable fixed points of mean-field model [5, 19]. For a discrete time dynamical system, a fixed point is called stable if it attracts all the trajectories that start from some neighborhood of the fixed point. Otherwise, a fixed point is unstable. If $f(x)$ is continuously differentiable in an open neighborhood of a fixed point x_0, a sufficient condition for x_0 to be stable or unstable is $|f'(x_0)| < 1$ or $|f'(x_0)| > 1$, respectively; see, e.g., [13].

Let $\text{Bin}(n,p)$ be a binomial random variable. Then the density ρ_t in the mean-field model satisfies the following stochastic recursion. Recall that $\deg(v)$ denotes the degree with respect to the long edges only, so the total degree of a vertex v is $\deg(v) + 4$.

Lemma 2. *For the mean-field approximation in a torus with N^2 nodes, ρ_t is a Markov process given by*

$$(4.1) \qquad N^2 \rho_{t+1} = \text{Bin}(N^2 \rho_t, f^+(\rho_t)) + \text{Bin}(N^2(1 - \rho_t), f^-(\rho_t)),$$

where

$$(4.2) \qquad f^+(x) = \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \sum_{i=k}^{n+1} \binom{n}{i-1} x^{i-1}(1-x)^{n-i+1},$$

$$(4.3) \qquad f^-(x) = \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \sum_{i=k}^{n} \binom{n}{i} x^i(1-x)^{n-i}.$$

Moreover, given ρ_t, ρ_{t+1} has mean $f(\rho_t)$ and variance $g(\rho_t)/N^2$ where

$$(4.4) \qquad f(x) = x f^+(x) + (1-x) f^-(x),$$

$$(4.5) \qquad g(x) = xf^+(x)(1 - f^+(x)) + (1 - x)f^-(x)(1 - f^+(x)).$$

Proof. Clear, since in the MF approximation, each vertex is assumed to have $\deg(v) + 4$ neighbors, each active with probability ρ_t, independently of each other and of $\deg(v)$; furthermore, different vertices are regarded as independent. $\qquad \square$

Remark 1. *In our model, the activation of a vertex is deterministic given the number of active vertices in the closed neighborhood. More generally, one can consider a model where an active [inactive] vertex with i active neighbors is activated with some probability p_i^+ $[p_i^-]$ (where $p_i^{\pm}$ are some given probabilities). In this more general case, (4.2)–(4.3) become*

$$(4.6) \qquad f^+(x) = \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \sum_{i=1}^{n+1} p_i^+ \binom{n}{i-1} x^{i-1}(1-x)^{n-i+1},$$

$$(4.7) \qquad f^-(x) = \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \sum_{i=0}^{n} p_i^- \binom{n}{i} x^{i}(1-x)^{n-i}.$$

Lemma 2 shows that the conditional variance of ρ_{t+1} is $g(\rho_t)/N^2 = O(N^{-2})$; thus ρ_{t+1} is well concentrated for large N, and we can approximate ρ_{t+1} by the mean $f(\rho_t)$.

The function $f(\cdot)$ given by (4.4) can be simplified to

$$(4.8) \qquad f(x) = xf^+(x) + (1-x)f^-(x)$$

$$= \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \sum_{i=k}^{n+1} \binom{n}{i-1} x^{i}(1-x)^{n-i+1}$$

$$+ \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \sum_{i=k}^{n} \binom{n}{i} x^{i}(1-x)^{n-i+1}$$

$$= \sum_{n=4}^{N^2-1} \mathbb{P}\left(\deg(v) = n - 4\right) \left(\sum_{i=k}^{n+1} \binom{n+1}{i} x^{i}(1-x)^{n-i+1} \right).$$

This can also be seen directly, since if $\deg(v) = n - 4$ with respect to the long edges, then the closed neigborhood of v contains $n + 1$ vertices, of which k have to be active for activation of v, and in the MF approximation, these $n + 1$ vertices are active independently of each other.

In Section 2.1 we showed that the degree distribution can be approximated by Poisson $\mathrm{Po}(\lambda)$ distribution. We use the last fact to approximate $f(x)$. Consider the function

$$(4.9) \qquad \bar{f}(x) = \bar{f}_k(x) = \sum_{n=4}^{\infty} \frac{e^{-\lambda}\lambda^{n-4}}{(n-4)!} \sum_{i=k}^{n+1} \binom{n+1}{i} x^{i}(1-x)^{n-i+1}.$$

The difference between $f(x)$ and $\bar{f}(x)$ can be bounded by

$$(4.10)$$

$$|f(x) - \bar{f}(x)| \leq \sum_{n=4}^{\infty} \left| \mathbb{P}\left(\deg(v) = n - 4\right) - \frac{e^{-\lambda}\lambda^{n-4}}{(n-4)!} \right| \sum_{i=k}^{n+1} \binom{n+1}{i} x^{i}(1-x)^{n-i+1}$$

$$\leq \sum_{n=4}^{\infty} \left| \mathbb{P}\left(\deg(v) = n - 4\right) - \frac{e^{-\lambda}\lambda^{n-4}}{(n-4)!} \right| = O\left(\frac{1}{N}\right)$$

where the last equality follows from Lemma 1.

4.2. **Derivation of criticality for various k values.** We rewrite $\bar{f} = \bar{f}_k$ defined in (4.9) as

$$(4.11) \qquad \bar{f}_k(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(\sum_{i=k}^{n+5} \binom{n+5}{i} x^i (1-x)^{n+5-i} \right).$$

For $k = 0, 1, 2$ and 3, respectively, the internal sums in Eq. (4.11) are

$$(4.12) \qquad \sum_{i=0}^{n+5} \binom{n+5}{i} x^i (1-x)^{n+5-i} = 1,$$

$$(4.13) \qquad \sum_{i=1}^{n+5} \binom{n+5}{i} x^i (1-x)^{n+5-i} = 1 - (1-x)^{n+5},$$

$$(4.14) \qquad \sum_{i=2}^{n+5} \binom{n+5}{i} x^i (1-x)^{n+5-i} = 1 - (1-x)^{n+5} - (n+5)(1-x)^{n+4}x,$$

$$(4.15) \qquad \sum_{i=3}^{n+5} \binom{n+5}{i} x^i (1-x)^{n+5-i} = 1 - (1-x)^{n+5} - (n+5)(1-x)^{n+4}x$$

$$- \frac{(n+5)(n+4)}{2}(1-x)^{n+3}x^2.$$

Hence, (4.11) yields, by simple calculations,

$$(4.16) \qquad \bar{f}_0(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} = 1,$$

$$(4.17) \qquad \bar{f}_1(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(1 - (1-x)^{n+5}\right) = 1 - e^{-\lambda x}(1-x)^5,$$

$$(4.18) \qquad \bar{f}_2(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(1 - (1-x)^{n+5} - (n+5)(1-x)^{n+4}x\right)$$

$$= 1 - e^{-\lambda x}\left((1-x)^5 + 5x(1-x)^4 + \lambda x(1-x)^5\right),$$

$$(4.19) \qquad \bar{f}_3(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \left(1 - (1-x)^{n+5} - (n+5)(1-x)^{n+4}x\right.$$

$$\left. - \frac{(n+5)(n+4)}{2}(1-x)^{n+3}x^2\right)$$

$$= 1 - e^{-\lambda x}\left((1-x)^5 + 5x(1-x)^4 + \lambda x(1-x)^5 + \frac{\lambda^2}{2}x^2(1-x)^5\right.$$

$$\left. + 5\lambda x^2(1-x)^4 + 10x^2(1-x)^3\right).$$

Proposition 1. *Let $\bar{f}_k(x) : [0,1] \to [0,1]$ be the family of maps for $k = 0, 1, 2, 3$ defined by (4.16)–(4.19). These maps have the following fixed points, for any $\lambda > 0$:*

(i) *for $k = 0$ the only fixed point is 1 and it is stable.*

(ii) *for $k = 1$ there are two fixed points: 1 is stable and 0 is unstable.*

(iii) *for $k = 2$ there are three fixed points: 0 and 1 are stable and $x_2(\lambda) \in (0, 1)$ is unstable;*

(iv) *for $k = 3$ there are three fixed points: 0 and 1 are stable and $x_3(\lambda) \in (0, 1)$ is unstable.*

Proof. For $k = 0$, the equation $\bar{f}_0(x) = x$ reduces to just

$$(4.20) \qquad\qquad x = 1.$$

In this case the fixed point $x = 1$ is stable since $\bar{f}_0'(x) = 0$.

For $k = 1$, $\bar{f}_1(x) = x$ can be written

$$(4.21) \qquad\qquad (1 - x)e^{\lambda x} = (1 - x)^5.$$

This equation has only two solutions 0 and 1 in $[0, 1]$, where 0 is an unstable fixed point since $\bar{f}_1'(0) = 5 + \lambda > 1$, while 1 is a stable fixed point because $\bar{f}_1'(1) = 0$.

For $k = 2$, $\bar{f}_2(x) = x$ we obtain

$$(4.22) \qquad\qquad (1 - x)e^{\lambda x} = (1 - x)^4(1 + 4x + \lambda x - \lambda x^2).$$

Clearly, 0 and 1 are solutions of (4.22). Divide both sides of (4.22) by $1 - x$ and let $g(x) = e^{\lambda x}$ be the LHS and $h(x)$ the RHS of the resulting equation.

Since $g(0) = h(0) = 1$, $g(1) = e^\lambda > 0 = h(1)$ and $h'(0) = 1 + \lambda > \lambda = g'(0)$, $h(x) = g(x)$ has a solution in $(0, 1)$.

This solutions is unique. To see this, first observe that $h''(x) = 0$ has a unique solution $x_{\text{infl}} \in (0, 0.5)$. Indeed, $h''(x)$ is a polynomial of degree three, i.e., it has at most three real roots. Clearly, $h''(1) = 0$, and since $h''(0) = -8\lambda - 18 < 0$ and $h''(0.5) = 0.5\lambda + 3 > 0$, by the intermediate value theorem $h''(x) = 0$ has a solution in $(0, 0.5)$. Moreover, $h'''(1) = -30 < 0$ and the leading coefficient of $h''(x)$ is positive; therefore, the third solution has to be greater than one.

Also, observe, that $h'(x) = 0$ has a unique solution $x_{\max} \in (0, 1)$. Indeed, $h'(x)$ is a polynomial of degree four, i.e., it has at most four real roots. Since $h'(1) = h''(1) = 0$ and $h'''(1) < 0$, 1 is a root of multiplicity two of $h'(x) = 0$ and $h'(1 \pm \varepsilon(\lambda)) < 0$ for $\varepsilon(\lambda)$ sufficiently small. Since the leading coefficient of $h'(x)$ is positive, $h'(x)$ must have two additional real roots, such that one is bigger and the other is smaller than 1. Moreover, $h'(0) > 0$, therefore, there is a unique root $x_{\max}$ in $(0, 1)$ and $h(x)$ has a unique maximum in $x_{\max}$ on $(0, 1)$.

Since $h(x)$ is concave on $[0, x_{\text{infl}})$ but $g(x)$ is convex, by Rolle's theorem they may have at most two intersections over that interval and one of those is at $x = 0$; furthermore, if there is an intersection in $(0, x_{\text{infl}})$, then $h(x_{\text{infl}}) < g(x_{\text{infl}})$.

On $(x_{\text{infl}}, 1)$, $h''(x) > 0$ and since also $h'(1) = 0$, it follows that $h'(x) < 0$ and $h(x)$ is decreasing on $(x_{\text{infl}}, 1)$. Furthermore, $g(x)$ is increasing, and thus the functions $h(x)$ and $g(x)$ may intersect at most once in $(x_{\text{infl}}, 1)$; moreover, if there is such an intersection, then $h(x_{\text{infl}}) > g(x_{\text{infl}})$.

Consequently, $h(x) = g(x)$ has a unique root $x_2(\lambda)$ in $(0, 1)$, and thus $x_2(\lambda)$ is the unique fixed point of $\bar{f}_2(x)$ in $(0, 1)$. In this case 0 and 1 are stable fixed points since $\bar{f}_2'(x)$ is zero at those points. The function $\bar{f}_2(x)$ is increasing on $[0, 1]$ and since the fixed points 0 and 1 are stable, one can see that $x_2(\lambda)$ is an unstable fixed point.

For $k = 3$, $\bar{f}_3(x) = x$ reduces to

$$(4.23) \quad (1-x)e^{\lambda x} = \frac{1}{2}(1-x)^3(2+(6+2\lambda)x+x^2(12+6\lambda+\lambda^2)-x^3(8\lambda+2\lambda^2)+\lambda^2x^4).$$

We note that Eq. (4.23) has solutions 0 and 1 for any λ.

There is also at least one solution $x_3(\lambda) \in (0,1)$ for any fixed $\lambda \geq 0$. Indeed, divide each side of Eq. (4.23) by $(1-x)$ and denote the LHS and the RHS of the resulting expression by $g(x) = e^{\lambda x}$ and $h(x)$, respectively. The fact that $g(0) = h(0) = 1$, $g(1) = e^{\lambda} > 0 = h(1)$, and $h'(0) = 1 + \lambda > \lambda = g'(0)$ verifies the statement.

Thus $g(x) = h(x)$ has at least two solutions in $[0,1)$, including $x = 0$, so by Rolle's theorem $g'(x) = h'(x)$ has at least one solution in $(0,1)$. If $g(x) = h(x)$ has more than one solution in $(0,1)$, then $g'(x) = h'(x)$ has by Rolle's theorem at least two solutions in $(0,1)$. Hence, it is enough to show that $g'(x) = h'(x)$ has a unique solution in $[0,1]$ to guarantee that the solution $x_3(\lambda)$ of $g(x) = h(x)$ in $(0,1)$ is unique.

The first two derivatives of $h(x)$ are

$$(4.24) \qquad h'(x) = 3\lambda^2 x^5 - (10\lambda^2 + 20\lambda)x^4 + (12\lambda^2 + 44\lambda + 24)x^3$$
$$- (6\lambda^2 + 27\lambda + 27)x^2 + (\lambda^2 + 2\lambda + 2)x + \lambda + 1$$

and

$$(4.25) \qquad h''(x) = 15\lambda^2 x^4 - (40\lambda^2 + 80\lambda)x^3 + (36\lambda^2 + 132\lambda + 72)x^2$$
$$- (12\lambda^2 + 54\lambda + 54)x + \lambda^2 + 2\lambda + 2.$$

First we show that $h''(x) = 0$ has two solutions on $[0,1]$, noting that 0 and 1 are not solutions since $h''(0) = \lambda^2 + 2\lambda + 2 > 0$ and $h''(1) = 20$. If $p(x)$ is a univariate polynomial of degree n then the variation of signs $V_c(p)$ of $p(x)$ at $x = c$ is the number of sign changes between consecutive elements in the sequence $\{p^{(j)}(x)\}_{j=0}^n \big|_{x=c} = \{p(c), p'(c), \ldots, p^{(n)}(c)\}$ (ignoring any terms that are 0). By the Budan–Fourier theorem, the number of roots on $[0,1]$ of $h''(x) = 0$ is $V_0(h'') - V_1(h'') - 2k$ where $k \in \mathbb{N} \cup \{0\}$. We have at $x = 0$

$$(4.26) \qquad \begin{aligned} h''(0) &= \lambda^2 + 2\lambda + 2 \\ h'''(0) &= -12\lambda^2 - 54\lambda - 54 \\ h^{(4)}(0) &= 72\lambda^2 + 264\lambda + 144 \\ h^{(5)}(0) &= -240\lambda^2 - 480\lambda \\ h^{(6)}(0) &= 360\lambda^2 \end{aligned}$$

and at $x = 1$

$$(4.27) \qquad \begin{aligned} h''(1) &= 20 \\ h'''(1) &= -30\lambda + 90 \\ h^{(4)}(1) &= 12\lambda^2 - 216\lambda + 144 \\ h^{(5)}(1) &= 120\lambda^2 - 480\lambda \\ h^{(6)}(1) &= 360\lambda^2 \end{aligned}$$

Clearly, $V_0(h'') = 4$ for any $\lambda > 0$. The derivatives $h'''(1)$, $h^{(4)}(1)$ and $h^{(5)}(1)$, as functions of $\lambda > 0$, change sign at $\lambda = 3$, $9 \pm \sqrt{69}$ and 4, respectively. It is easily verified by inspection, considering the intervals $(0, 9 - \sqrt{69}]$, $(9 - \sqrt{69}, 3]$, $(3, 4]$, $(4, 9 + \sqrt{69}]$ and $(9 + \sqrt{69}, \infty)$ separately, that $V_1(h'') = 2$ for every $\lambda > 0$.

Therefore, the number of solutions on $[0,1]$ to $h''(x) = 0$, counted with multiplicity, is at most $V_0(h'') - V_1(h'') = 2$. Furthermore, by (4.26) and (4.27) $h''(0) > 0$ and $h''(1) > 0$ for any $\lambda > 0$. However, $h''(0.2) = -0.256\lambda^2 - 4.16\lambda - 5.92 < 0$ for any $\lambda > 0$. Thus, $h''(x) = 0$ has exactly two roots x_1 and x_2 in $[0,1]$, and $0 < x_1 < 0.2 < x_2 < 1$.

Consider now the third derivative of the function $h(x)$,

$$(4.28)\quad h'''(x) = 60\lambda^2 x^3 - 120\lambda(\lambda+2)x^2 + (72\lambda^2 + 264\lambda + 144)x - 12\lambda^2 - 54\lambda - 54.$$

For $x = 0.2$,

$$\begin{aligned}
h'''(0.2) &= -1.92\lambda^2 - 10.8\lambda - 25.2 \\
h^{(4)}(0.2) &= 31.2\lambda^2 + 168\lambda + 144 \\
h^{(5)}(0.2) &= -168\lambda^2 - 480\lambda \\
h^{(6)}(0.2) &= 360\lambda^2
\end{aligned}$$

These alternate in sign, for any $\lambda > 0$. Thus, by a Taylor expansion at $x = 0.2$, $h'''(x) < 0$ for $x \leq 0.2$.

Consider first the interval $[0, 0.2]$. Since $h''' < 0$ there, h' is concave, and since g' is convex, $h' - g'$ is concave. Furthermore, $h'(0) - g'(0) = 1 > 0$. It follows that $h' - g'$ has at most one root in $[0, 0.2]$, and that if there is such a root, then $h'(0.2) \leq g'(0.2)$.

Next, on the interval $[0.2, x_2) \subset (x_1, x_2)$, we have $h''(x) < 0$ and thus h' is decreasing. Since g' is increasing, $h' - g'$ has at most one root in $(0.2, x_2)$, and if there is such a root, then $h'(0.2) > g'(0.2)$.

Finally, in the interval $[x_2, 1]$, $h'' \geq 0$ and thus h' is increasing and $h'(x) \leq h'(1) = 0$. Since $g' > 0$, there is no root of $h' = g'$ in $[x_2, 1]$.

Combining these cases, we see that there is at most one root of $g'(x) = h'(x)$ in $[0,1]$. As said above, this implies that there is a unique root $x_3(\lambda)$ in $(0,1)$ of $g(x) = h(x)$, and thus a unique fixpoint $x_3(\lambda)$ in $(0,1)$ of $\bar{f}_3$.

One can check that $\bar{f}_3'(x) = 0$ at 0 and 1, and thus these two fixed points are stable. Furthermore, $\bar{f}_3(x)$ is increasing on $[0,1]$ and using the stability of the fixed points 0 and 1, and the uniqueness of the fixed point $x_3(\lambda)$, it follows that $x_3(\lambda)$ is an unstable fixed point. $\qquad\square$

For all cases considered above 0 is a fixed point of $\bar{f}$. As we noted before, the error $f(x) - \bar{f}(x)$ is 0 at 0, so this fixed point is also a fixed point of $f(x)$ for any N. If x is an unstable fixed point of $\bar{f}$ with $\bar{f}'(x) > 1$, then (4.10) implies that $f(x)$ has a fixed point shifted from x at most by $O(1/N)$. These arguments are valid in case λ is a fixed constant independent of N.

Let p denote the probability that a node is initially activated and p_c be the nontrivial solution(s) derived above. Since ρ_t is a Markov process, for the mean-field approximation we obtain the following theorem.

Theorem 2. *In the mean-field approximation of the activation process $A(t)$ over random graph $G_{\mathbb{Z}_N^2, p_d}$ there exists a critical probability p_c such that for a fixed p, with high probability for large N, all vertices will eventually be active if $p > p_c$, while all vertices will eventually be inactive for $p < p_c$. The value of p_c is given as the function of k and λ as follows:*

    (i) *For $k = 0$ and any λ, $p_c = 0$ and all vertices will become active in one step for any p.*

(ii) *For $k = 1$ and any λ, $p_c = 0$, i.e., for any fixed $p > 0$, all vertices will eventually become active with high probability.*

(iii) *For $k = 2$ and any λ, $p_c = x_2(\lambda)$, where $x_2(\lambda) \in (0, 1)$ is a nontrivial solution to $x = \bar{f}_2(x)$.*

(iv) *For $k = 3$ and any λ, $p_c = x_3(\lambda)$, where $x_3(\lambda) \in (0, 1)$ is a nontrivial solution to $x = \bar{f}_3(x)$.*

Proof. Consider the case $0 \le p < p_c$ (and thus (iii) or (iv)); the case $p_c < p \le 1$ is similar and (i) and (ii) are trivial. In the limit as $N \to \infty$, $\rho_0 = p$ and ρ_t is deterministic with $\rho_{t+1} = \bar{f}(\rho_t)$. Since $p < p_c$, the sequence $\rho_t = \bar{f}^t(p)$ converges, as $t \to \infty$, to the fixpoint 0. Furthermore, because $\bar{f}'(0) = 0$, the convergence is (at least) quadratic, and in particular geometric.

Now consider a fixed positive integer N. The deterministic sequence $\bar{f}^t(p)$ just considered reaches below $1/N$ for $t \ge t_N$, where $t_N = O(\log N)$. The sequence ρ_t is a random perturbation of $\bar{f}^t(p)$. In each step, we have two sources of error: the difference in mean $f(\rho_t) - \bar{f}(\rho_t) = O(1/N)$, by (4.10), and the random error coming from the binomial distributions in (4.1), which by a standard Chernoff bound is $O(N^{-0.9})$ with probability $1 - O(N^{-1})$, say. Since further $|f'(x)| \le 1$ for small x, the combined error from the first t_N steps is $t_N(O(N^{-1}) + O(N^{-0.9})) = O(N^{-0.8})$ with probability $1 - O(t_N N^{-1}) = 1 - o(1)$. Hence, with high probability, we reach a state with $\rho_t = O(N^{-0.8})$. Then $f(\rho_t) = O(\rho_t^2) = O(N^{-1.6})$, and by another Chernoff bound (or Chebyshev's inequality), $\rho_{t+1} = O(N^{-1.6})$ with high probability. But then $f(\rho_{t+1}) = O(\rho_{t+1}^2) = O(N^{-3.2})$, and thus (conditionally given ρ_{t+1}), the expected number of active vertices at time $t + 2$ is $N^2 f(\rho_{t+1}) = O(N^{-1.2}) = o(1)$, and thus with high probability there are no active vertices at all at time $t + 2$. $\square$

Lemma 3. *For $k = 2, 3$, $p_c = x_k(\lambda)$, is a non-increasing function of $\lambda \ge 0$.*

Proof. If we increase λ, then the average number of edges is increased. Moreover, if $G_N(\lambda)$ denotes the random graph $G_{\mathbb{Z}_N^2, p_d}$ with parameter λ, and $\lambda_1 < \lambda_2$, then we can couple the random graphs $G_N(\lambda_1)$ and $G_N(\lambda_2)$ such that $G_N(\lambda_1) \subseteq G_N(\lambda_2)$; it is then evident that if all vertices eventually are activated in $G_N(\lambda_1)$ (for a given initially active set), then so are all vertices in $G_N(\lambda_2)$. The same holds for the mean-field approximation, where again we can couple two models with parameters λ_1 and λ_2, with $\lambda_1 < \lambda_2$, such that the set of activated vertices for λ_1 is a subset of the set of activated vertices for λ_2. It follows that in Theorem 2, $p_c(\lambda_1) \ge p_c(\lambda_2)$, i.e., p_c is a non-increasing function of λ.

An alternative, analytical, proof is given in the appendix. $\square$

Corollary 1. *Cases (iii) and (iv) of Theorem 2 can be sharpened as follows.*

(iii) *For $k = 2$ and any λ, $p_c = x_2(\lambda)$, where $x_2(\lambda) \in (0, x_2(0)]$ is a unique solution to $x = \bar{f}_2(x)$ and $x_2(0) \approx 0.132$.*

(iv) *For $k = 3$ and any λ, $p_c = x_3(\lambda)$, where $x_3(\lambda) \in (0, x_3(0)]$ is a unique solution to $x = \bar{f}_3(x)$ and $x_3(0) = 0.5$.*

Proof. The values $x_2(0) = \frac{11}{12} - \frac{1}{12}(235 + 6\sqrt{1473})^{1/3} - \frac{13}{12}(235 + 6\sqrt{1473})^{-1/3} \approx 0.131123$, and $x_3(0) = \frac{1}{2}$ can be easily obtained from (4.22) and (4.23), respectively. By Lemma 3, for $k = 2, 3$ we have that $x_k(\lambda) \le x_k(0)$, i.e., Corollary 1 holds. $\square$

It is also easy to see that for any fixed $p > 0$, if λ is large enough, then the proportion of vertices active after the first step is more than, say, $0.6 > p_c$, and thus eventually all vertices will be active. Consequently, $p_c = x_k(\lambda) \to 0$ as $\lambda \to \infty$.

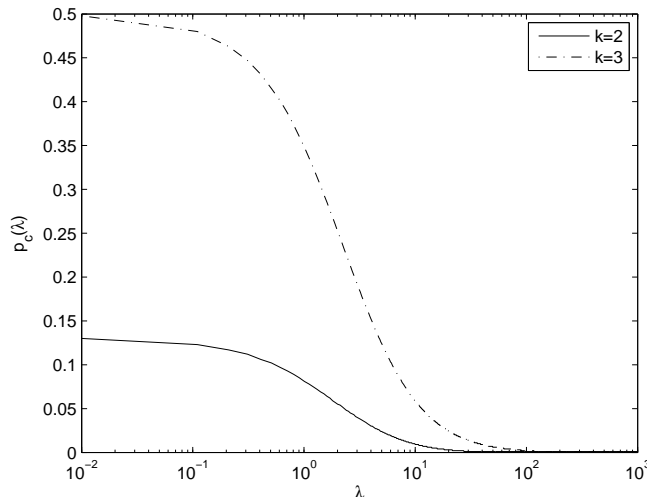FIGURE 1. p_c as a function of λ for $k = 2, 3$.

In Theorem 2 we obtained p_c with respect to λ for different k values. Notice, that in the case of $\lambda \to 0$, the solution tends to 0.5 and to 0.131123 for $k = 3$ and $k = 2$, respectively. The dependence of p_c on λ (for $k = 2, 3$) is shown on Figure 1.

Remark 2. *Note that for $k \geq 5$, an inactive vertex will remain inactive for ever. Hence, unless all vertices are activated at the beginning, there is at each step a set of inactive vertices. Furthermore, every neighbor of an inactive vertex becomes inactive; hence, for the graph $G_{\mathbb{Z}_N^2, p_d}$, every vertex will become inactive after at most N steps. For the mean-field approximation, every vertex has at least a fixed positive probability of becoming inactive at every step; hence (almost surely) every vertex will eventually become inactive in the mean-field approximation, too.*

5. Concluding remarks

In this work we introduced the random graph model $G_{\mathbb{Z}_N^2, p_d}$. We derived bounds on the diameter of this graph and described its degree distribution. We studied the activation processes on $G_{\mathbb{Z}_N^2, p_d}$ in mean-field approximation. Specifically, we derived conditions for phase transitions as a function of initialization probability p and long edge parameter λ. Moreover, we showed the boundedness and monotonicity of the critical probability as the function of λ. One can also see from Figure 1 that p_c drops significantly for $\lambda \in (0.1, 10)$.

Open questions include the generalization of these results for other lattice types and higher dimensions. An ongoing research concerns activity propagation in $G_{\mathbb{Z}_N^2, p_d}$ with two types of nodes (excitatory and inhibitory) [16].

Mean-field results provide useful insight on the activation processes. Nevertheless, an important future task remains to investigate the activation process without using the "averaging" feature of the mean-field approach. Open problems concerning the properties of $G_{\mathbb{Z}_N^2, p_d}$ and the activation process on the graph include: What is the number of small cycles? What is the clustering coefficient? A rough $O(\log N)$

upper bound on time required to activate all vertices in supercritical case follows from the proof of Theorem 2. It would be interesting to obtain a sharper estimate.

References

[1] Aizenman, M., Kesten, H., and Newman, C.M, *Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation*, Commun. Math. Phys., (1988).

[2] Balister, P., Bollobás, B., and Kozma, R., *Large deviations for mean field models of probabilistic cellular automata*, Random Structures & Algorithms **29, (3)**, 399-415, (2006).

[3] Barbour, A.D., Holst, L., and Janson, S., *Poisson Approximation*, Clarendon Press, Oxford, (1992).

[4] Benjamini, I., and Berger, N., *The diameter of long-range percolation clusters on finite cycles*, Random Structures & Algorithms, **19 (2)**, 102-111, (2001).

[5] Biskup, M., Chayes, L., and Crawford, N., *Mean-Field Driven First-Order Phase Transitions in Systems with Long-Range Interactions*, J. of Statistical Physics, **122 (6)**, 1139-1193, (2006).

[6] Bollobás, B., and Chung, F. R. K., *The Diameter of a Cycle Plus a Random Matching*, SIAM J. Disc. Math., **1 (3)**, 328-333, (1988).

[7] Bollobás, B., Janson, S., and Riordan, O., *The phase transition in inhomogeneous random graphs*, Random Structures & Algorithms, **31 (1)**, 3-122, (2007).

[8] Chung, F., and Lu, L., *The Diameter of Sparse Random Graphs*, Advances in Applied Mathematics, **26, (4)**, 256–279, (2001).

[9] Coppersmith, D., Gamarnik, D., and Sviridenko, M., *The diameter of a long-range percolation graph*, Random Structures & Algorithms, **21, (1)**, 1-13, (2002).

[10] Erdős, P., and Rényi, A., *On the evolution of random graphs*, Magyar Tudoamányos Akadémia, Mat. Kut. Int. Közl., **5**, 17-61, (1960).

[11] Freeman, W.J., *Mechanism and significance of global coherence in scalp EEG*, Current Opinion in Neurobiology, **31**, 199-205, (2015).

[12] Freeman, W.J., *The Physiology of Perception*, Scientific American, **264**, 78-85, (1991).

[13] Hirsch, M. W., Smale, S., and Devaney, R. L., *Differential equations, dynamical systems, and an introduction to chaos*. Academic Press, (2012).

[14] Janson, S., Łuczak, T., Turova, T., and Vallier, T., *Bootstrap percolation on the random graph $G_{N,p}$*, The Annals of Applied Probability, **22 (5)**, 1989-2047, (2012).

[15] Kozma, R., and Puljic, M., *Random graph theory and neuropercolation for modeling brain oscillations at criticality*, Current opinion in neurobiology, **31**, 181-188, (2015).

[16] Kozma, R., Ruszinkó, M., and Sokolov, Y., *Percolation on a power-law-like random graph coupled with a lattice. Part II: The case of two types of nodes*, in progress.

[17] Le Cam, L., *An approximation theorem for the Poisson binomial distribution*, Pacific Journal of Mathematics, **10**, 1181-1197, (1960).

[18] Newman, M. E. J., and Watts, D. J., *Scaling and percolation in the small-world network model*, Phys. Rev. E, **60**, 7332 - 7342 (1999).

[19] Talagrand, M., *Mean Field Models for Spin Glasses*, Vol. 1 & 2, Springer-Verlag Berlin Heidelberg, (2011).

[20] Turova, T., and Vallier, T., *Bootstrap percolation on a graph with random and local connections.*, arXiv:1502.01490, (2015).

[21] Watts, D. J., and Strogatz, S. H., *Collective dynamics of 'small-world' networks*, Nature, 440-442 (1998).

6. Appendix. Analytic proof of Lemma 3

Proof. For $k = 2$ or 3, denote the functions $g(x)$ and $h(x)$ from corresponding cases of the proof of Proposition 1 by $g_k(x)$ and $h_k(x)$, and let $F_k(\lambda, x) = g_k(x) - h_k(x)$. Then $x_k(\lambda)$ is a root of $F_k(\lambda, x) = 0$. We have shown that $g_k(x) = h_k(x)$ has a unique root in $(0, 1)$ and the proofs also show that the root is simple, so $\partial F_k/\partial x = g_k'(x) - h_k'(x) \neq 0$ at $x = x_k(\lambda)$. It follows from the implicit function theorem that

$x_k(\lambda)$ is an infinitely differentiable function of $\lambda \in (0, \infty)$, and that

$$(6.1) \qquad \frac{dx_k(\lambda)}{d\lambda} = -\frac{\partial F_k/\partial \lambda}{\partial F_k/\partial x}(\lambda, x_k(\lambda)).$$

Now, $F_k(\lambda, 1) = g_k(1) - h_k(1) > 0$, and thus $F_k(\lambda, x) > 0$ for $x > x_k(\lambda)$. Consequently, $\partial F_k/\partial x > 0$ at $x = x_k(\lambda)$.

In the case $k = 2$, the numerator of (6.1) is

$$(6.2) \qquad \frac{\partial F_2}{\partial \lambda}(\lambda, x_2(\lambda)) = x_2(\lambda)\left(e^{\lambda x_2(\lambda)} - (1 - x_2(\lambda))^4\right) > 0,$$

because $(1 - x_2(\lambda))^4 < 1 \le e^{\lambda x_2(\lambda)}$. Hence, $dx_2(\lambda)/d\lambda < 0$ for any $\lambda > 0$.

For $k = 3$, the numerator of (6.1) is given by

$$(6.3) \qquad \frac{\partial F}{\partial \lambda}(\lambda, x) = x\left(e^{\lambda x} - (1 - x)^3(1 + 4x + \lambda x - \lambda x^2)\right) = xF_2(\lambda, x)$$

at $x = x_3(\lambda)$. We claim that $h_3(x) > h_2(x)$ for $x \in (0, 1)$. Since $g_2(x) = g_3(x)$, this implies $F_3(\lambda, x) < F_2(\lambda, x)$. In particular, $F_2(\lambda, x_3(\lambda)) > F_3(\lambda, x_3(\lambda)) = 0$ and thus (6.3) yields $\partial F/\partial \lambda(\lambda, x_3(\lambda)) > 0$; hence $dx_3(\lambda)/d\lambda < 0$ by (6.1).

To verify the claim, we calculate $h_3(x) - h_2(x) = \frac{1}{2}x^2(1 - x)^2(\lambda^2 x^2 - (2\lambda^2 + 10\lambda)x + \lambda^2 + 10\lambda + 20)$. The square factors are positive on $(0, 1)$ and the quadratic factor has roots $1 + (5 \pm \sqrt{5})/\lambda$. Since $1 + (5 \pm \sqrt{5})/\lambda > 1$ for any $\lambda > 0$, and furthermore the quadratic factor is positive at $-\infty$, it is positive on $(-\infty, 1)$. Hence $h_3(x) - h_2(x) > 0$ on $(0, 1)$. $\qquad \square$

(Svante Janson) Department of Mathematics, Uppsala University,
PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: svante.janson@math.uu.se

(Robert Kozma) Department of Mathematics, The University of Memphis,
Memphis, TN, 38152, USA
E-mail address: rkozma@memphis.edu

(Miklós Ruszinkó) Alféd Rényi Institute of Mathematics,
Hungarian Academy of Sciences,
13-15 Reáltanoda utca, Budapest, Hungary, 1053
E-mail address: ruszinko.miklos@renyi.mta.hu

(Yury Sokolov) Department of Mathematics, The University of Memphis,
Memphis, TN, 38152, USA
E-mail address: ysokolov@memphis.edu