Bootstrap percolation on the random regular graph

József Balogh* and Boris G. Pittel †

December 8, 2005

Dedicated to Alan Frieze on the occasion of his 60*-th birthday.*

Abstract

The k-parameter bootstrap percolation on a graph is a model of an interacting particle system, which can also be viewed as a variant of a growth process of a cellular automata with threshold $k \geq 2$. At the start each of the graph vertices is *active* with probability p and *inactive* with probability $1-p$, independently of other vertices. Presence of active vertices triggers a percolation process controlled by the recursive rule: an active vertex remains active forever, and a currently inactive vertex becomes active when at least k of its neighbors are active. The basic problem is to identify, for a given graph, p^-, p^+ such that for $p < p^-$ ($p > p^+$ resp.) the probability that all vertices are eventually active is very close to 0 (1 resp.). The percolation process is a Markov chain on the space of subsets of the vertex set, which is easy to describe but hard to analyze rigorously in general. We study the percolation on the random d-regular graph, $d \geq 3$, via analysis of the process on its multigraph counterpart. Here, thanks to a "principle of deferred decisions", the percolation dynamics is described by a surprisingly simple Markov chain. Its generic state is formed by the counts of

*University of Illinois at Urbana-Champaign, work was partially done while at the Ohio State University; email: jobal@math.uiuc.edu, research supported in part by NSF grant DMS-0302804 and OTKA grant 049398.

†The Ohio State University; email: bgp@math.ohio-state.edu, research supported in part by NSF grant DMS-0406024.

currently active and nonactive vertices having various degrees of percolation capabilities. We replace the chain by a deterministic dynamic system, and use its integrals to show – via exponential supermartingales – that the percolation process undergoes relatively small fluctuations around the deterministic trajectory. This allows us to show existence of the phase transition within an interval $[p^-(n), p^+(n)]$, such that (1) $p^{\pm}(n) \to p^* = 1 - \min_{y \in (0,1)} y/\mathbb{P}\left(\mathrm{Bin}(d-1, 1-y) < k\right)$; (2) $p^+(n) - p^-(n)$ is of order $n^{-1/2}$ for $k < d-1$, and $n^{-\varepsilon_n}$, ($\varepsilon_n \downarrow 0, \varepsilon_n \log n \to \infty$), for $k = d-1$.

1 Introduction

In a seminal paper [25] McCullogh and Pitts attempted to describe the brain as a network of interacting neurons. The neurons are visualized as the vertices of a directed graph with weighted edges. At each moment a vertex can be in one of two states, 0 ("passive") and 1 ("active"). If the weighted sum of the states of a vertex' neighbors exceeds (resp. falls below) a given threshold value, the vertex state at the next moment is 1 (resp. 0). Given an initial data, this switching rule determines, in principle, the time evolution of the configuration of active vertices. A "local" nature of the rule has made a computer simulation of the dynamic networks a very efficient tool. McCullogh and Pitts proved that this scheme is broad enough to include a universal Turing machine as a particular case. The neuron network model is mathematically close to the cellular automata introduced later by Ulam [28] and von Neumann see [14]. A popular cellular automaton is Conway's 'Game of Life' (see [20] and [9], Ch. 19). Unlike the neuron networks, the underlying graph of a cellular automaton is usually the two-dimensional grid, thus has a rigid built-in structure.

As often with appealing mathematical models, similar dynamic processes have been suggested in other venues, such as spread of epidemics, gossips, percolation, voter models, interacting particles etc. (See, for instance, Durrett [17], Grimmett [21] and Liggett [24].)

In this paper we study a dynamic version of percolation known as a *bootstrap* percolation. Let G be a connected graph. Assume that each vertex can be in one of two possible states, *inactive* or *active*. At time 0, each vertex is active with probability p, independently of all other vertices. At time t, where t is a positive integer, the state of each vertex is updated: if a vertex was active at time $t-1$ then it stays active; otherwise it will be activated if it had at least k active neighbors at time $t-1$. Here $k \geq 2$ is given. Introduce $P(G,p)$ the probability that each vertex sooner or later will become active. It is intuitively obvious, and can be proved rigorously, that $P(G,p)$ is a nondecreasing function of p. (The case $k=1$ is excluded, since here $P(G,p)$ is simply the probability that at least one vertex is initially active, which is $1-(1-p)^{|V(G)|}$.) For a constant $0 \leq c < 1$ let

$$p_c = \inf\{p : P(G,p) > c\}.$$

Given (small) positive constant ε one would like to determine the ε-*window* of the process, i.e. to find the interval $[p_\varepsilon, p_{1-\varepsilon}]$.

The bootstrap percolation was first mentioned and studied –in statistical physics context– by Chalupa, Reich and Leath [16] for the d-regular infinite tree T_d. To be sure, they considered a "dual" process: given an integer m, at each round every active vertex with fewer than m active vertices is deleted. However, for $k = d - m + 1$ the two processes are equivalent. Very recently, Balogh, Peres and Pete [7] studied a broader class of the infinite trees and graphs. Their general estimates enabled them to obtain an alternative proof of the result in [16]. In the language of the k-parameter process, $P(T_d, p) = 0$ or 1 dependent on whether p is less or more than p^*, the supremum of all p for which the equation

$$\mathbb{P}\big(\mathrm{Binom}(d-1,(1-x)(1-p)) \leq d-k-1\big) = x \tag{1}$$

has a real root $x \in (0,1)$. Furthermore,

$$p^*|_{k=d-1} = 1 - \frac{1}{d-1} \qquad \text{and} \qquad p^*|_{k=2} = 1 - \frac{(d-2)^{2d-5}}{(d-1)^{d-2}(d-3)^{d-3}}. \tag{2}$$

p^* can be interpreted as a critical probability.

After pioneering work [16], van Enter [18] Schonmann [27] studied the bootstrap percolation on the infinite 2-dimensional grid and the d-dimensional grid respectively. In particular, Schonmann proved that if $k \leq d$ then the critical probability of the process is 0, otherwise 1.

A substantial effort has been put into analysis of the bootstrap process for the "mixed" case, i. e. for G belonging to a sequence $\{G_n\}_{n\geq 1}$ of ever-larger finite graphs G_n. Popular cases are the $n \times n$ square grid or tori, the n-dimensional hypercube, etc. In this case p is allowed to depend on n, and the central issue is an asymptotic behavior of $p_\varepsilon(n)$ and $p_{1-\varepsilon}(n)$. Is $p_{1-\varepsilon}(n) - p_\varepsilon(n)$ going to zero as $n \to \infty$, and if so, is $\lim_{n\to\infty} p_{1-\varepsilon}(n)/p_\varepsilon(n)$ equal to 1, or at least bounded? "Yes" means intuitively that there is a sharp phase transition in the limiting behavior of the percolation process. And if there is a phase transition, does the transition interval shrink to a limiting point, or no limit exists? If the limiting point exists and non-zero, then it is usually called a *critical probability.*

Aizenmann and Lebowitz [1] found the bounds for $p_c(n)$ in the case of the d-dimensional grid $[1, n]^d$ and $k = 2$. Cerf and Manzo [15] generalized the bounds for parameter $2 \le k \le d$. The central case $k = d = 2$ was also studied in [4], and the culmination point was reached by Holroyd [22], who discovered a sharp asymptotic formula for the critical probability. Balogh and Bollobás [6] determined up to constant factor $p_c(n)$ for the case of the n-dimensional cube and $k = 2$.

The question whether the transition is sharp, i. e. whether $\lim_{n\to\infty} p_{1-\varepsilon}(n)/p_\varepsilon(n) = 1$ for all $\varepsilon > 0$ remained wide open in most of the cases. As a notable exception, Balogh and Bollobás [5] proved that for $[1, n]^d$ and $k = 2$, the transition is indeed sharp. (The proof utilizes a powerful result of Friedgut and Kalai [19] on sharp thresholds for monotone graph properties.) It is probably true that the transition is sharp for all $k \in [2, d]$, but no rigorous proof has been found.

That the sharp results have been so elusive is due to analytical complexity of a Markov chain describing the dynamics of percolation process on the underlying graph. Our goal in this paper is a detailed study of the case when the graph G_n is the random d-regular graph $G_{n,d}$. (For the pioneering work on $G_{n,d}$ see Bender and Canfield [8], Bollobás [10], [11], Wormald [29], [30].) Since in the vicinity of each vertex w.h.p. the graph $G_{n,d}$ looks like a subtree of the infinite tree T_d, it seems plausible that p^* (see (2)) is the critical probability for $G_{n,d}$. We are not aware of any existing proof of this deceptively simple conjecture. The issue of the width of phase transition interval $[p_\varepsilon(n), p_{1-\varepsilon}(n)]$ is even murkier, as it has no natural counterpart for the percolation on T_d. The only related result we know of is due to Bollobás [13], who proved that the critical probability for $G_{n,d}$ is bounded away from 0 for all $2 \le k < d$.

We use a multigraph version of $G_{n,d}$, known as "configuration model"

([10], [11]) and a "principle of deferred decisions" to describe the dynamics of percolation as a Markov chain, which is perfectly tailored for asymptotic study. (The substitution is justified since the random multigraph is in fact simple with probability bounded away from zero as $n \to \infty$.) A generic state is the counts of currently active and nonactive vertices of various degrees of percolation capabilities. We replace the chain by a deterministic dynamic system, and use its integrals to show – via exponential supermartingales – that the percolation process undergoes relatively small fluctuations around the deterministic trajectory. The techniques are reminiscent of those used by Pittel, Spencer and Wormald [26] for a k-core problem, by Aronson, Frieze and Pittel [3] for a maximum matching problem on random graphs, and by Aldous and Pittel [2] for a study of random graphs with immigrating vertices. The approximation estimates are sharp enough to identify the critical probability, and thus to confirm the conjecture, and also to estimate the critical width. The bound of the latter turns out to be dependent on whether $k < d-1$ or $k = d-1$.

2 Main results

First we introduce some definitions and notation. Let $G_{n,d}$, the random d-regular graph on n vertices, with $d \geq 3$ be the underlying graph of the bootstrap percolation process. (It is well-known, [29], that $G_{n,d}$ is connected with probability $1 - o(1)$, $n \to \infty$.) Assume that, for each vertex, the excitation threshold is $k \geq 2$. Suppose there is given a *deterministic* subset $\mathcal{A}$ of vertices that are active at time $t = 0$; $\mathcal{I} = \mathcal{A}^c$ is the set inactive vertices at time $t = 0$. Let p denote $\mathcal{A}$'s density, i. e. $p = A/n$, $A = |\mathcal{N}|$. Let $\mathcal{I}_f$ denote the random final set of inactive vertices for the bootstrap percolation process initiated by the active set $\mathcal{A}$, and $I_f = |\mathcal{I}_f|$. The complete activation means that $I_f = 0$. The distribution of I_f depends, of course, on p. (For $k = 1$ the limiting probability of complete activation is 1 even if $A = 1$, or $p = 1/n$.) Our task is to identify the critical density and to bound the transition window enclosing it.

Introduce

$$p^* = 1 - \inf_{y \in (0,1)} R(y,p), \quad R(y,p) \overset{def}{=} \frac{y}{\mathbb{P}\left(\mathrm{Bin}(d-1, 1-y) < k\right)},$$

where $\mathrm{Bin}(d-1, 1-y)$ stands for the binomially distributed random variable, with parameters $d-1$ and $1-y$. For $k < d-1$ the infimum is attained at an interior point y^*; in particular for $k = 2$ and $d > 3$,

$$y^* = \frac{(d-1)(d-3)}{(d-2)^2}, \quad p^* = 1 - \frac{(d-2)^{2d-5}}{(d-1)^{d-2}(d-3)^{d-3}}.$$

For $k = d - 1$,

$$p^* = 1 - F(0+) = 1 - (d-1)^{-1}.$$

By the definition of p^*, the equation

$$p = 1 - R(y, p)$$

has no roots $y \in (0, 1)$ if $p > p^*$. We will prove that for $p < p^*$ this equation has exactly one root $\hat{y}(p)$ in $(0, 1)$ if $k = d - 1$, and exactly two roots in $(0, 1)$ if $k < d-1$, in which case it is the larger root that we denote by $\hat{y}(p)$.

Suppose that $2 \leq k < d - 1$.

Theorem 1. *Let* $\omega = \omega(n) \to \infty$ *slowly enough, so that* $\log \omega = o(\log n)$. *(i) If* $p > p^* + \omega n^{-1/2}$, *then*

$$\lim_{n\to\infty} \mathbb{P}(I_f = 0) = 1.$$

(ii) If $p \leq p^* - \omega n^{-1/2}$, *then for each* $\gamma > 1/2$

$$\lim_{n\to\infty} \mathbb{P}(I_f = nh(\hat{y}(p)) + O\big(n^{1/2}\omega^{\gamma}(p^* - p)^{-1/2}\big)) = 1,$$

where

$$h(y) \stackrel{def}{=} (1-p)\mathbb{P}(Bin(d, 1-y) < k).$$

Note. Thus, for $k < d - 1$, p^* is the critical probability for complete activation. The part (ii) means intuitively that for p noticeably below p^*, the number of eventually active neighbors of an initially inactive vertex is distributed binomially in the limit, with success probability equal $1 - \hat{y}(p)$. The estimates show that the transition window has width of order $n^{-1/2}$ at most.

Suppose that $k = d - 1$. In this case if there is a cycle formed by the initially inactive vertices then these vertices will remain inactive since at no time any of them will have more than $d-2$ active neighbors. Thus there will be no complete activation. Now, with the positive limiting probability, $G_{n,d}$ does contain short cycles, of length 3 for example, [30]. Therefore, for every fixed density $p<1$, with a positive limiting probability there is a cycle of length 3, with all three vertices inactive at $t=0$. Hence $\limsup_n \mathbb{P}(I_f = 0) < 1$ for every such p. In other words, the limiting probability of complete activation is below 1 for every $p<1$. If $I_f \neq 0$, its natural to consider the density $n^{-1}I_f$ of the inactive "diehards". We say that w.h.p. there is an *almost* complete activation if $n^{-1}I_f \to_P 0$.

Theorem 2. *(i) Suppose $p-p^* \geq n^{-\varepsilon}$, where $\varepsilon = \varepsilon(n) \downarrow 0$ and $\varepsilon \log n \to \infty$. Then*

$$\lim_{n\to\infty} \mathbb{P}(I_f = O((p-p^*)^{-3/2})) = 1.$$

(ii) Let $p \leq p^* - \omega n^{-\sigma}$, $\sigma = \frac{1}{2(d+5)}$, $\omega = \omega(n) \to \infty$, $\log \omega = o(\log n)$. Then

$$\lim_{n\to\infty} \mathbb{P}(I_f = nh(\hat{y}(p)) + O(n^{1-3\sigma}\omega^{6\sigma}(p^*-p)^{-1})) = 1,$$

and the remainder term is negligible compared with the term $nh(\hat{y}(p))$.

Note. Thus, for $k=d-1$, p^* is the critical probability for almost complete activation. If the bounds are tight, and we think they are, then the transition window enclosing p^* is quite asymmetric, unlike the case $k<d-1$. While its lower part remains polynomially short, its length decreases as $n^{-\frac{1}{2(d+5)}}$, which far exceeds $n^{-1/2}$, the total window width for $k<d-1$. The upper part of the window shrinks to p^* as well, but infinitely slower than n^{-a} for any fixed $a>0$.

Suppose that the set $\mathcal{A}$ of the initially active vertices is random, and independent of $G_{n,d}$. Suppose that $A = |\mathcal{A}|$ satisfies a condition

$$\lim_{n\to\infty} \mathbb{P}(n^{-1}|A - \mathbb{E}[A]| \leq \lambda n^{-1/2}) = 1,$$

for any $\lambda = \lambda(n) \to \infty$, $A = \mathbb{E}[A] + O_P(n^{1/2})$ in short. These conditions are met when, for instance, each vertex is initially active with probability p, independently of all other vertices *and* $G_{n,d}$.

Theorem 3. *Under the above conditions, the statements in Theorems 1 and 2 hold for the parameter $p \overset{def}{=} n^{-1}\mathbb{E}[A]$.*

Here is why. *Conditioned* on the set $\mathcal{A}$, the random graph is still distributed as $G_{n,d}$, whence we have our standard case with a deterministic set of initially active vertices. Thus Theorems 1, 2 apply. For ω defined in Theorem 1, pick $\lambda \to \infty$ such that $\lambda = o(\omega)$. We know that

$$\mathbb{P}(B_n) \to 1, \quad B_n \overset{def}{=} \{|A - \mathbb{E}[A]| \le \lambda n^{1/2}\}.$$

It remains to observe that on the event B_n we can replace $n^{-1}A$ by its expected value $p = n^{-1}\mathbb{E}[A]$. Indeed in all cases covered in Theorems 1, 2 we have $|p - p^*| \ge \omega n^{-1/2}$, and on the event B_n

$$|n^{-1}\mathbb{E}[A] - n^{-1}A| \le \lambda n^{-1/2} \ll \omega n^{-1/2}.$$

□

3 Bootstrap percolation process on $G_{n,d}$.

3.1 Markov Chain

We begin with a model of percolation process on a given graph $G = (V, E)$ in which interactions between vertices are organized in rounds. We start with a partition $V = \mathcal{A}(0) \biguplus \mathcal{I}(0)$, $\mathcal{A}(0)$ and $\mathcal{I}(0)$ being interpreted as the initial set of active vertices and the initial set of inactive vertices. Recursively, given a partition $V = \mathcal{A}(t) \biguplus \mathcal{I}(t)$ after $t \ge 0$ rounds, we set

$$\begin{aligned} \mathcal{A}(t+1) &= \mathcal{A}(t) \cup \{v \in \mathcal{I}(t) : |\Gamma(v) \cap \mathcal{A}(t)| \ge k\}, \\ \mathcal{I}(t+1) &= V \setminus \mathcal{A}(t+1); \end{aligned}$$

here $\Gamma(v)$ is the set of neighbors of v in G. In words, every vertex still inactive after t rounds that has at least k active neighbors becomes active after the $(t+1)$-th round. The process terminates once either $\mathcal{A}(t) = \emptyset$ or $\mathcal{I}(t) = \emptyset$.

Consider the case when G is $G_{n,d}$, the (uniformly) random d-regular graph on $V = [n]$. Clearly the distribution of $\{\mathcal{A}(t)\}_{t\ge 0}$ depends only on $|\mathcal{A}(0)|$. Further, at each round we need to know only the neighborhoods of currently inactive vertices, and partitions of those neighborhoods into active and non-active vertices. Conditioned on the information about $G_{n,d}$ gained after several rounds, $G_{n,d}$ remains uniformly distributed on the set of all d-regular graphs compatible with this information. The power of these observations is severely limited though, and an asymptotic study of the process is highly problematic. To better appreciate the inherent difficulties the reader may wish to look, for instance, at just the distribution of $|\mathcal{A}(1)|$.

Our task is to find a surrogate percolation process which we can study and be able to use to prove Theorems 1, 2. One important simplification is that instead of rounds of activation we will use pairwise interactions, between an active vertex and its inactive neighbor. (This is certainly analogous to epidemic models in which it is customary to view pairwise interactions as driving force in spreading epidemic diseases.) Each such interaction increases activation level of an inactive vertex by 1. And once at some moment this level reaches k, the inactive vertex becomes active. For this "slowed down" percolation process the time is measured in those interactions, and at each moment we have a partition of the inactive vertex set into k, some possibly empty, subsets of vertices with activation levels $0, 1, \dots, k-1$. Of course, the terminal set of inactive vertices for this process is the same as for the original one. Since each step results in a relatively small change of the graph state, one can expect that w.h.p. the scaled realization of percolation process will be close to a solution of the approximating system of differential equations. This device is quite similar to the approach developed in [26] for a core problem.

However, pairwise interaction between two vertices can occur at most once: $G_{n,d}$ is a simple graph, so parallel edges are forbidden. Because of this restriction, the Markov chain for the localized percolation process on $G_{n,d}$ is forbiddingly complex. Our second crucial simplification is that instead of $G_{n,d}$ we consider $MG_{n,d}$, the random d-regular *multi*graph. (In [3] a similar substitution was used for analysis of a matching algorithm on a sparse random graph discovered by Karp and Sipser [23].) Now the parallel edges, and even the loops, are allowed! (d-regularity means that for each vertex v the double number of loops at v plus the number of parallel edges incident to v is d.) $MG_{n,d}$ is generated as follows, Bollobás [11]. Introduce sets $S_1, \dots, S_n$, $|S_i| \equiv d$, representing the vertices $1, \dots, n$, respectively. Assuming that nd is even, there are $(nd-1)!!$ complete matchings on $\mathcal{S} = \cup_i S_i$. Let $\mathcal{M}_n$ be the uniformly random matching on S. If $s' \in S_i, s'' \in S_j$ are partners in $\mathcal{M}_n$ we draw an edge between vertices i and j. If $i = j$, then we draw a loop at vertex i. The resulting random multigraph is $MG_{n,d}$, as it is uniformly distributed on the set of all d-regular multigraphs. As we will see shortly, the percolation process on $MG_{n,d}$ is perfectly amenable to asymptotic study. The question is then how do we translate the results in terms of $G_{n,d}$ itself?

The key is that, conditioned on the event $\mathcal{S}_n = \{\text{no loops, no parallel edges}\}$, $MG_{n,d} \stackrel{\mathcal{D}}{\equiv} G_{n,d}$. Furthermore

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{S}_n) = \exp(-(d-1)/2 - (d-1)^2/4) > 0,$$

[10], [30]. So, for every multigraph property Q_n,

$$\mathbb{P}(G_{n,d} \in Q_n) = O(\mathbb{P}(MG_{n,d} \in Q_n)),$$

whence

$$\mathbb{P}(G_{n,d} \in Q_n) = o(1) \quad \text{if} \quad \mathbb{P}(MG_{n,d} \in Q_n) = o(1).$$

We are ready to describe the bootstrap percolation on $MG_{n,d}$, or rather on the random matching $\mathcal{M}_n$ which determines $MG_{n,d}$. Let $\mathcal{A}(0) \subset [n]$ be a set of vertices active at $t = 0$. Let $\mathcal{I}(0) = [n] \setminus \mathcal{A}(0)$ denote the set of inactive vertices at $t = 0$. It is natural to call all points of a subset S_i active (inactive) if the "host" vertex i is active (inactive). For the first step, we pick a pair $\{s_i, s_j\} \in \mathcal{M}_n$, $s_i \in S_i$, $s_j \in S_j$ such that, $i \in \mathcal{A}(0)$, i. e. s_i is active, and then delete both s_i and s_j from S_i and S_j respectively. Recursively, after t steps we will have $\mathcal{A}(t)$, $\mathcal{I}(t)$, the sets of currently active and inactive vertices, and the sets $S_i(t) \subseteq S_i$, $i \in [n]$. At step $t+1$ we (a) pick an active point

$$s' \in S_i(t), \quad i \in \mathcal{A}(t),$$

(b) identify its partner

$$s'' \in S_j(t), \quad j \in [n],$$

(c) delete both s' and s'' from the sets $S_i(t)$ and $S_j(t)$. If s'' is currently inactive, i. e. hosted by a currently inactive vertex i, and s'' is the k-th point deleted from the initial set $S_i(0) = S_i$, then i and the remaining points in $S_i(t+1)$ are declared active from this moment on.

To specify the process completely, we label the points of S by indices from 1 to nd, and at each step we pick as s' an active point with the lowest index.

Given the information on the active points and their partners picked and deleted in the first t steps, the random matching on the $nd - 2t$ points still present remains uniformly distributed. Effectively it means that, conditioned on the full prehistory, at step $t+1$ we select s'' uniformly at random in the set $\bigcup_{i\in[n]} S_i(t) \setminus \{s'\}$.

Let $\mathbf{S}(t) = \{S_i(t)\}_{i\in[n]}$. The process $\{\mathbf{S}(t), \mathcal{A}(t), \mathcal{I}(t)\}_{t\geq 0}$ is obviously Markov. Introduce

$$I_j(t) = |\{i \in \mathcal{I}(t) : |S_i(t)| = d - j\}|, \quad 0 \leq j \leq k-1;$$

$I_j(t)$ is the total count of currently inactive vertices whose points were picked as partners of active points j times after t steps. In particular, $I_{k-1}(t)$ is the total count of those vertices which are one pairwise interaction away from becoming active. Introduce

$$I(t) = \sum_{j=0}^{k-1} (d-j) I_j(t),$$

the total count of currently inactive points, and likewise

$$A(t) = \sum_{i\in\mathcal{A}(t)} |S_i(t)|,$$

the total count of currently active points. Let $\mathbf{I}(t) = (I_0(t), \dots, I_{k-1}(t))$. It is easy to see that $\{A(t), \mathbf{I}(t)\}_{t\geq 0}$ is also Markov. Let us compute the one-step transition probabilities. There are three possibilities.
(i) s'' is currently active. The (conditional) probability of this outcome is

$$\frac{A(t)-1}{A(t)-1+I(t)}. \tag{3}$$

The next state is

$$\begin{aligned} A(t+1) &= A(t) - 2, \\ I_j(t+1) &= I_j(t), \ (0 \leq j \leq k-1). \end{aligned} \tag{4}$$

(ii) s'' is in a currently inactive set $S_i(t)$ of cardinality $d-j$, with $j < k-1$. The probability of this outcome is

$$\frac{(d-j)I_j(t)}{A(t)-1+I(t)}. \tag{5}$$

The next state is

$$\begin{aligned} A(t+1) &= A(t) - 1, \\ I_\ell(t+1) &= I_\ell(t), \quad (\ell \neq j, j+1), \\ I_j(t+1) &= I_j(t) - 1, \\ I_{j+1}(t+1) &= I_{j+1}(t) + 1. \end{aligned} \tag{6}$$

(iii) s'' is in a currently inactive set $S_i(t)$ of cardinality $d - k + 1$. The probability of this outcome

$$\frac{(d-k+1)I_{k-1}(t)}{A(t)-1+I(t)}. \tag{7}$$

The next state is

$$\begin{aligned} A(t+1) &= A(t) + d - k - 1, \\ I_j(t+1) &= I_j(t), \quad j < k-1, \\ I_{k-1}(t+1) &= I_{k-1}(t) - 1. \end{aligned} \tag{8}$$

Averaging over the possible transitions, we obtain the equations for expectations $\mathbb{E}[\cdot|\circ]$ of $A(t{+}1), I_0(t{+}1), \ldots, I_{k-1}(t{+}1)$ conditioned on $A(t), I_0(t), \ldots, I_{k-1}(t)$:

$$\begin{aligned} \mathbb{E}[A(t+1)|\circ] =& A(t) + \frac{(d-k+1)(d-k)I_{k-1}(t) - 2A(t) - I(t) + 2}{A(t)+I(t)-1}, \\ \mathbb{E}[I_0(t+1)|\circ] =& I_0(t) - \frac{dI_0(t)}{A(t)+I(t)-1}, \\ \mathbb{E}[I_j(t+1)|\circ] =& I_j(t) + \frac{(d-j+1)I_{j-1}(t) - (d-j)I_j(t)}{A(t)+I(t)-1}, \quad (0 < j \leq k-1). \end{aligned} \tag{9}$$

3.2 Differential equations approximation

The equations (9) make it plausible that $n^{-1}A(n\tau)$, $n^{-1}I_j(n\tau)$ are likely to be close to $a(\tau)$, $i_j(\tau)$, the (deterministic) solution of the system of differential equations

$$\begin{aligned} a' &= \frac{[(d-k+1)(d-k)]i_{k-1} - 2a - i}{a+i}, \\ i_0' &= \frac{-di_0}{a+i}, \\ i_j' &= \frac{(d-j+1)i_{j-1} - (d-j)i_j}{a+i}, \quad (0 < j \leq k-1), \end{aligned} \tag{10}$$

under the initial conditions

$$a(0) = dp,\ i(0) = d(1-p),\ i_0(0) = 1-p,\ i_j(0) = 0,\ (0 < j \leq k-1). \tag{11}$$

The system (10) has a surprisingly simple solution. Since $A(t) + I(t) = A(0) + I(0) - 2t$, we should expect that

$$a(\tau) + i(\tau) = a(0) + i(0) - 2\tau.$$

This is indeed true, because it follows from (10) directly that

$$\frac{d}{d\tau}(a + i) = -2.$$

Introduce

$$u = u(\tau) = \ln(a(\tau) + i(\tau))^{-1/2}, \tag{12}$$

so that $\tau = 0$ corresponds to $u_0 := u(0) = \ln(a(0) + i(0))^{-1/2}$. Using (12) we obtain *linear* equations for $f_j(u) = i_j(u(\tau))$, $(0 \le j \le k-1)$:

$$\begin{aligned} \frac{df_0}{du} &= -\, df_0, \\ \frac{df_j}{du} &=(d-j+1)f_{j-1} - (d-j)f_j, \quad j \ge 1. \end{aligned} \tag{13}$$

For arbitrary u_1 and initial values $f_i(u_1)$, the solution is

$$f_j(u) = e^{-(d-j)(u-u_1)} \sum_{r=0}^{j} \binom{d-r}{j-r} (1 - e^{-(u-u_1)})^{j-r} f_r(u_1), \tag{14}$$

$0 \le j \le k-1$. This formula is obviously true for $j = 0$. Suppose it holds for some $j \ge 0$. Then using

$$\frac{d}{du}(f_{j+1}e^{(d-j-1)(u-u_1)}) = e^{(d-j-1)(u-u_1)}(d-j)f_j(u),$$

and the inductive hypothesis, we obtain

$$\begin{aligned} &f_{j+1}(u)e^{(d-j-1)(u-u_1)} - f_{j+1}(u_1) \\ &\quad = (d-j)\sum_{r=0}^{j} f_r(u_1)\binom{d-r}{j-r} \int_{u_1}^{u} (1 - e^{-(v-u_1)})^{j-r} e^{-(v-u_1)}\, dv \\ &\quad = \sum_{r=0}^{j} f_r(u_1)\frac{d-j}{j-r+1}\binom{d-r}{j-r}(1 - e^{-(u-u_1)})^{j-r+1} \\ &\quad = \sum_{r=0}^{j} f_r(u_1)\binom{d-r}{j+1-r}(1 - e^{-(u-u_1)})^{j+1-r}. \end{aligned}$$

And this proves (14) for $f_{j+1}(u)$.
Going back to τ, for $u_1 = u_0$ we transform (14) into

$$i_j(\tau) = y^{d-j}\sum_{r=0}^{j} i_r(0)\binom{d-r}{j-r}(1-y)^{j-r}, \quad y = \left(\frac{a(\tau)+i(\tau)}{a(0)+i(0)}\right)^{1/2}. \qquad (15)$$

In particular, for the conditions $a(0) = dp$, $i_0(0) = 1-p$, $i_j(0) = 0$, $j > 0$, we have then

$$i_j(\tau) = (1-p)\binom{d}{j}y^{d-j}(1-y)^j, \quad y = (1-2\tau/d)^{1/2}.$$

Consequently

$$\begin{aligned} a(\tau) =& d - 2\tau - \sum_{j=0}^{k-1}(d-j)i_j(\tau) = dS(y,p), && (16)\\ S(y,p) :=& y^2 - (1-p)yP(y) = yP(y)[R(y)-(1-p)], && (17)\\ P(y) :=& \mathbb{P}(\mathrm{Bin}(d-1, 1-y) < k), && (18)\\ R(y) :=& \frac{y}{P(y)}. && (19) \end{aligned}$$

From the origin of $S(y,p)$ it is clear that our focus is on $\hat{y}(p) = \inf\{y : S(y,p) > 0\}$. Conjecturally, the percolation process terminates around the time $t \sim n\hat{\tau}(p)$, $\hat{\tau}(p)$ being determined by

$$\hat{y}(p) = (1-2\hat{\tau}(p)/d)^{1/2}.$$

According to (17), we need then to have a close look at $R(y)$ defined in (19). We have

$$R'(y) = \frac{P(y) - yP'(y)}{P^2(y)}.$$

Using the definition (18) of $P(y)$, after some algebra we obtain

$$P'(y) = (d-k)\binom{d-1}{k-1}y^{d-k-1}(1-y)^{k-1}. \qquad (20)$$

Therefore

$$P(y) - yP'(y) = -(d-1-k)\binom{d-1}{k-1}y^{d-k}(1+O(y)) < 0, \quad y \downarrow 0,$$

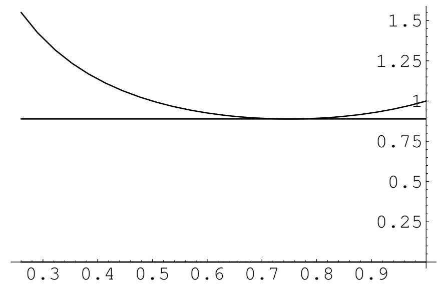

Figure 1: The function $R(y)$ for $d = 4,\ k = 2$, with infinum 8/9.

if $k < d-1$, because

$$P(y) = \binom{d-1}{k-1} y^{d-k}(1 + O(y)), \quad y \downarrow 0.$$

At the other end,

$$P(1) - P'(1) = P(1) = 1 > 0.$$

So, staying with the case $k < d-1$, $\inf\{R(y) : y \in (0,1]\}$ is attained at some point of $(0,1)$. Since $P(y) - yP'(y) < 0$ (resp. > 0) for $y \downarrow 0$ (resp. $y \uparrow 1$), and

$$\begin{aligned}&(P(y) - yP'(y))' = -yP'' \\ =& -(d-k)\binom{d-1}{k-1} y^{d-k-1}(1-y)^{k-2}[(d-k-1) - y(d-2)], \quad y \in (0,1),\end{aligned}$$

$P(y) - yP'(y) = 0$ has a unique root $y^* \in (0,1)$. (Otherwise there would be at least three roots, counting multiplicities, and $P'' = 0$ would have at least two interior roots.) In particular, a simple calculation shows that for $k = 2$, and $d \geq 4$,

$$y^* = \frac{(d-1)(d-3)}{(d-2)^2}, \quad R(y^*) = \frac{(d-2)^{2d-5}}{(d-1)^{d-2}(d-3)^{d-3}}.$$

Using the definition (17) of $S(y,p)$ and $k < d-1$, we see that $S(y,p) \sim y^2 > 0$, $y \downarrow 0$, uniformly for p. From the above discussion it follows that in fact $S(y,p) > 0$ for all $y \in (0,1]$, iff

$$p > p^* \stackrel{\text{def}}{=} 1 - R(y^*) = 1 - \frac{(d-2)^{2d-5}}{(d-1)^{d-2}(d-3)^{d-3}}.$$

For $p = p^*$, the graph of $S(y,p)$ touches the interval $[0,1]$ from above at the unique positive point y^*; in particular

$$S(y^*,p^*) = 0, \quad S_y(y^*,p^*) = 0, \quad S_{yy}(y^*,p^*) \geq 0.$$

Let $p < p^*$. Since $R(y)$ decreases for $y < y^*$, and increases for $y > y*$, there are exactly two points in $(0,1)$ –$y_1(p) < y^*$ and $y_2(p) > y^*$– where the graph of $S(y,p)$ crosses y-axis; in particular, $y_2(p)$ is $\hat{y}(p)$ introduced above. And $y_1(p) \uparrow y^*$, $y_2(p) \downarrow y^*$ as $p \uparrow p^*$. Let $y(p)$ denote a minimum point of $S(y,p)$ for $p < p^*$. Clearly $y(p) \in (y_1(p), y_2(p))$, so that $y(p) \to y^*$ as $p \uparrow p^*$, and

$$S(y(p),p) < 0, \quad S_y(y(p),p) = 0, \quad S_{yy}(y(p),p) \geq 0.$$

We need to extend $y(p)$ to $p > p^*$, and to determine a sharp estimate of $S(y(p),p)$. To this end, let us prove first that $S_{yy}(y(p),p) > 0$ for $y \geq y(p)$ and $p \leq p^*$.

Since $S(y,p) \sim y^2$, $y \downarrow 0$, $S(y,p)$ has a (local) positive maximum on $(0,y(p))$. So $S_{yy}(y,p) = 0$ has at least two roots in $(0,y(p))$, separated by the maximum point. Furthermore, direct computation yields

$$S_{yy}(y,p) = 2 - (1-p)(d-k)\binom{d-1}{k-1} \times y^{d-1-k}(1-y)^{k-2}(d-k+1-dy).$$

It follows from this formula that $S_{yy}(y,p) > 0$ for $y \downarrow 0$ and $y \uparrow 1$. If $S_{yy}(y,p)$ changes its sign, or has a root of even multiplicity, somewhere between $y(p)$ and 1, then $S_{yy}(p) = 0$ has at least four roots, counting multiplicities. If so, $S_{yyy}(y,p)$ must vanish at three or more points in $(0,1)$. But, differentiating the last expression, we obtain that for $k \geq 3$

$$S_{yyy}(y) = by^{d-k-2}(1-y)^{k-3}(a_2y^2 + a_1y + a_0),$$

and for $k = 2$

$$S_{yyy}(y) = \beta_1 y^{d-4}(\alpha_1 y + \alpha_0).$$

So $S_{yyy}(y)$ may vanish only at two interior points. Thus $S_{yy}(y,p)$ is positive, and bounded away from 0, uniformly for $y \geq y(p)$, $p \leq p^*$.

Positivity of $S_{yy}(y,p)$ on $[y(p),1)$ implies that $y(p)$ is a unique stationary point of $S(y,p)$ on $[y_1(p),1)$. And $y(p)$ is continuously differentiable as a function of p on $[0,p^*]$. Moreover, $y(p)$, defined as a root of $S_y(y,p) = 0$, has

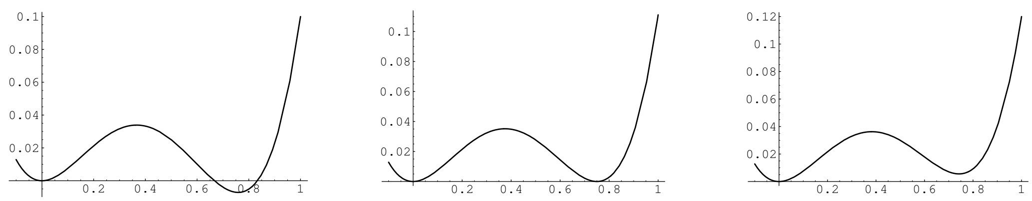

Figure 2: The function $S(y,p)$ for $d=4,\ k=2$ and $p<p^*,\ p=p^*,\ p>p^*$.

a differentiable extension for p close enough to p^* from above. This time $y(p)$ is the *local* minimum point of $S(y,p)$ on $(0,1]$. Thus, for $\varepsilon>0$ sufficiently small, $y(p)$ is well defined on $[0,p^*+\varepsilon]$ and

$$S_{yy}(y,p)\geq c_1>0,\quad (y\in[y(p),1],\quad p\in[0,p^*+\varepsilon]). \tag{21}$$

Convexity of $S(y,p)$ on $[y(p),1]$ implies that in fact $y(p)$ is the absolute minimum point of $S(y,p)$ on $[y(p),1]$. Finally, using differentiability of $y(p)$,

$$\begin{aligned}\frac{dS(y(p),p)}{dp}=&\left(\frac{\partial S}{\partial y}\frac{dy}{dp}+\frac{\partial S}{\partial p}\right)\bigg|_{y=y(p)}=\frac{\partial S}{\partial p}\bigg|_{y=y(p)}\\ =&y(p)\mathbb{P}(\mathrm{Bin}(d-1,1-y(p))<k)>0,\quad (\forall p\leq p^*+\varepsilon).\end{aligned}$$

Therefore, for $p\to p^*$,

$$S(y(p),p)\sim c(p-p^*),\quad c=S_{yy}(y^*,p^*)>0. \tag{22}$$

Consider now the case $2\leq k=d-1$. It is easy to see that in this case

$$S(y,p)=y^2-y(1-p)[1-(1-y)^{d-1}].$$

In particular 0 is the root of $S(y,p)=0$ of multiplicity 2. Furthermore

$$S(y,p)\geq y^2[1-(1-p)(d-1)],$$

so $S(y,p)>0,\ y\in(0,1]$, if $p>p^*$, where

$$p^*=1-(d-1)^{-1}.$$

If $p<p^*$, then

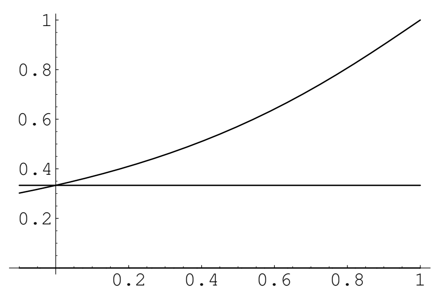

Figure 3: The function $R(y)$ for $d = 4,\ k = 3$ with infinum 1/3.

$$S(y,p) \sim \frac{y^2(p-p^*)}{1-p^*} < 0, \quad y \downarrow 0.$$

As $S(1,p) = p > 0$, $S(y,p) = 0$ has another root $\hat{y}(p) \in (0,1)$. From convexity of $(1-y)^{d-1}$ it follows that this root is unique. Unlike the case $k < d-1$, $\hat{y}(p) \downarrow 0$ as $p \uparrow p^*$. More precisely

$$\hat{y}(p) \sim \frac{2(d-1)}{d-2} \cdot (p^* - p), \quad p \uparrow p^*. \tag{23}$$

Thus we anticipate that the percolation process is likely to terminate after

$$\frac{nd}{2}(1-\hat{y}(p)^2) = \frac{nd}{2} - O(n(p^*-p)^2)$$

of pairwise interactions, i. e. after using almost all pairs in $\mathcal{M}_n$.

Furthermore, $S(y,p)$ attains its negative minimum at $y(p) \in (0, \hat{y}(p))$, and a simple computation yields: for $p \uparrow p^*$

$$y(p) \sim \frac{4(d-1)}{3(d-2)} \cdot (p^* - p), \tag{24}$$

$$S(y(p),p) \sim -c_1(p^*-p)^3. \tag{25}$$

In addition, for $p \uparrow p^*$,

$$S_{yy}(y(p),p) \sim c_2(p^*-p). \tag{26}$$

From

$$S_{yyy}(y,p) = (1-p)(d-1)(d-2)(1-y)^{d-4}(3-dy)$$

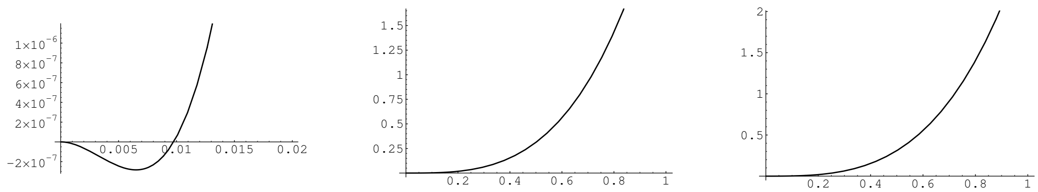

Figure 4: The function $S(y,p)$ for $d=4,\ k=3$ and $p<p^*,\ p=p^*,\ p>p^*$.

it follows, like in the case $k<d-1$, that $S_{yy}(y,p)>0$ for $y\ge y(p)$. In fact, for $y<d/3$, $S_{yyy}(y,p)>0$ and consequently

$$S_{yy}(y(p),p)\le S_{yy}(y,p)\le S_{yy}(y(p),p)+c_3(y-y(p)),\quad y\in[y(p),3/d].$$

Therefore, see (26),

$$S_y(y,p)\ge S_y(y(p),p)+S_{yy}(y(p),p)(y-y(p))\ge c_3(p^*-p)(y-y(p)). \quad (27)$$

On the other hand, invoking (25),

$$\begin{aligned} S(y,p)\le\ & S(y(p),p)+\tfrac{1}{2}S_{yy}(y(p),p)(y-y(p))^2+\tfrac{c_3}{6}(y-y(p))^3\\ \le\ & c_4(-(p^*-p)^3+(p^*-p)(y-y(p))^2+(y-y(p))^3). \end{aligned} \quad (28)$$

Both (27) and (28) hold for $y\in[y(p),3/d]$.

3.3 Concentration of the process

Our task now is to show that w.h.p. the percolation process stays close to the deterministic trajectory described in the previous section.

To this end we need to find some "good" functions $F_j(\mathbf{x})$, ($\mathbf{x}:=(a,i_0,\dots,i_{k-1})$), $0\le j\le k-1$, that remain constant on every solution of (10). Replacing in (14) u_1 with u, and u with 0 say, we obtain

$$e^{(d-j)u}\sum_{r=0}^{j}\binom{d-r}{j-r}(1-e^u)^{j-r}f_r(u)=f_j(0),\quad 0\le j\le k-1. \quad (29)$$

Given u, let $M(u)$ be a lower-triangular $k\times k$ matrix with

$$[M(u)]_{jr}=e^{(d-j)u}\binom{d-r}{j-r}(1-e^u)^{j-r},\quad 0\le r\le j\le k-1.$$

Comparison of (14) and (29) shows that actually

$$[M^{-1}(u)]_{jr} = e^{-(d-j)u}\binom{d-r}{j-r}(1-e^{-u})^{j-r}, \quad 0 \le r \le j \le k-1. \tag{30}$$

The entries of $M^{-1}(u)$ are uniformly continuous on $[0,\infty)$. Now $e^u = (a+i)^{-1/2}$, so

$$(a+i)^{-(d-j)/2}\sum_{r=0}^{j}\binom{d-r}{j-r}\left(1-(a+i)^{-1/2}\right)^{j-r} i_r \equiv \text{ const}, \quad 0 \le j \le k-1,$$

on every solution of (10). So we are led to the k functions

$$F_j(\mathbf{x}) = (a+i)^{-(d-j)/2}\sum_{r=0}^{j}\binom{d-r}{j-r}\left(1-(a+i)^{-1/2}\right)^{j-r} i_r,$$

$(0 \le j \le k-1)$, whose values are preserved on every solution $\mathbf{x}(\tau)$ of (10). Equivalently, denoting the right hand side vector function in (10) by $\mathbf{R}(\mathbf{x})$, we have

$$\mathbf{R}(\mathbf{x})^* \text{grad} F_j(\mathbf{x}) \equiv 0, \quad 0 \le j \le k-1. \tag{31}$$

Notice also that each $F_j(\mathbf{x})$ is of order $(a+i)^{-d/2}$ for $a+i \to 0$, and its first and second order derivatives are of order $(a+i)^{-(d+2)/2}$ and $(a+j)^{-(d+4)/2}$ respectively.

Denote $\mathbf{X}(t) = (A(t), I_0(t), \ldots, I_{k-1}(t))$. Following the method developed in [26], and in [3], we will use these integrals to show that w.h.p. $F_j(n^{-1}\mathbf{X}(t))$ stays very close to $F_j(n^{-1}\mathbf{X}(0))$, as long as $A(t)+I(t)$ remains sufficiently large.

Proposition 4. *Let $\omega = \omega(n) \to \infty$ however slowly. Introduce T the first time t when either* $\min\{A(t), I(t)\} = 0$ *or* $A(t)+I(t) < n\omega^{-\frac{1}{d+2}}$. *Then*

$$\mathbb{P}\left\{\max_{t \le T} |F_j(n^{-1}\mathbf{X}(t)) - F_j(n^{-1}\mathbf{X}(0))| > z\right\} = O\left(e^{-z(n/\omega)^{1/2}}\right),$$

uniformly for $z > 0$. Consequently, for some absolute constant $c > 0$,

$$\mathbb{P}\left\{\max_{t \le T} \left|n^{-1}\mathbf{X}(t) - \mathbf{x}(n^{-1}t)\right| \le z\right\} = 1 - O\left(e^{-cz(n/\omega)^{1/2}}\right).$$

Proof. Consider $j = 0$ for instance.
(I) Suppose that t is such that $\min\{A(t), I(t)\} > 0$ and $A(t)+I(t) \geq n\omega^{-\frac{1}{d+2}}$, i. e. $t < T$ in other words. Introduce

$$Q(t) = \exp\left\{(n/\omega)^{1/2}\left[F_0(n^{-1}\mathbf{X}(t)) - F_0(n^{-1}\mathbf{X}(0))\right]\right\}.$$

Let us evaluate $\mathbb{E}[Q(t+1)|\circ]$. First of all,

$$Q(t+1) = Q(t)\exp\left\{(n/\omega)^{1/2}\left[F_0(n^{-1}\mathbf{X}(t+1)) - F_0(n^{-1}\mathbf{X}(t))\right]\right\}. \quad (32)$$

Here

$$F_0(n^{-1}\mathbf{X}(t+1)) - F_0(n^{-1}\mathbf{X}(t)) = (n^{-1}\mathbf{X}(t+1) - n^{-1}\mathbf{X}(t))^*\mathrm{grad}F_0(\tilde{\mathbf{x}}), \quad (33)$$

where $\tilde{\mathbf{x}}$ is a point on the line segment connecting $n^{-1}\mathbf{X}(t+1)$ and $n^{-1}\mathbf{X}(t)$. Since $|\mathbf{X}(t+1) - \mathbf{X}(t)|$ is bounded by an absolute constant, we see that

$$A(t+1) + I(t+1) \geq 0.5n\omega^{-\frac{1}{d+2}}.$$

Consequently

$$|\mathrm{grad}F_0(\tilde{\mathbf{x}})| \leq c(\tilde{a} + \tilde{i})^{-(d+2)/2} \leq c_1\omega^{1/2}.$$

It follows then from (33) that

$$(n/\omega)^{1/2}|F_0(n^{-1}\mathbf{X}(t+1)) - F_0(n^{-1}\mathbf{X}(t)| \leq c_2 n^{-1/2}. \quad (34)$$

Therefore

$$\begin{aligned} &\exp\left\{(n/\omega)^{1/2}\left[F_0(n^{-1}\mathbf{X}(t+1)) - F_0(n^{-1}\mathbf{X}(t))\right]\right\} \quad (35) \\ &\leq 1 + (n/\omega)^{1/2}(F_0(n^{-1}\mathbf{X}(t+1)) - F_0(n^{-1}\mathbf{X}(t)) + c_3 n^{-1}. \end{aligned}$$

Applying to the increment of F_0 in this formula the second-order Taylor approximation yields: for some intermediate $\hat{\mathbf{x}}$ on the line segment $[n^{-1}\mathbf{X}(t+1), n^{-1}\mathbf{X}(t)]$,

$$\begin{aligned} &F_0(n^{-1}\mathbf{X}(t+1)) - F_0(n^{-1}\mathbf{X}(t)) \quad (36) \\ &= (n^{-1}\mathbf{X}(t+1) - n^{-1}\mathbf{X}(t))^*\mathrm{grad}F_0(n^{-1}\mathbf{X}(t)) + O(n^{-2}(\hat{a} + \hat{i})^{-(d+4)/2}). \end{aligned}$$

Since

$$\hat{a} + \hat{i} \geq 0.5\omega^{-\frac{1}{d+2}},$$

the remainder term in (36) is $O(n^{-1})$. So (35) reduces to

$$\begin{aligned}&\exp\left\{(n/\omega)^{1/2}\left[F_0(n^{-1}\mathbf{X}(t+1))-F_0(n^{-1}\mathbf{X}(t))\right]\right\} \\ &\quad\le 1+(n/\omega)^{1/2}(n^{-1}\mathbf{X}(t+1)-n^{-1}\mathbf{X}(t))^{*}\mathrm{grad}F_0(n^{-1}\mathbf{X}(t))+c_4n^{-1}.\end{aligned} \tag{37}$$

Since the right hand side in (37) is linear in $\mathbf{X}(t+1)$, we can compute its conditional expected value using (9). Denoting by $\mathcal{R}(\mathbf{X}(t))$ the differences $\mathbb{E}[A(t+1)|\circ]-A(t),\ldots,\mathbb{E}[I_{k-1}(t+1)|\circ]-I_{k-1}(t)$, determined in (9), and comparing them with $\mathbf{R}(\mathbf{x})$ (the right hand side of the differential equations (10)), we evaluate

$$\begin{aligned}&\mathbb{E}\left[(n^{-1}\mathbf{X}(t+1)-n^{-1}\mathbf{X}(t))^{*}\mathrm{grad}F_0(n^{-1}\mathbf{X}(t))|\circ\right] \\ =&n^{-1}\mathcal{R}(\mathbf{X}(t))^{*}\mathrm{grad}F_0(n^{-1}\mathbf{X}(t)) \\ =&n^{-1}\mathbf{R}(n^{-1}\mathbf{X}(t))\mathrm{grad}F_0(n^{-1}\mathbf{X}(t))+O(n^{-1}(A(t)+I(t))^{-1}).\end{aligned} \tag{38}$$

The remainder term follows from the observation that $|\mathcal{R}(\mathbf{X})|$ is bounded by a constant and that consequently

$$|\mathcal{R}(\mathbf{X}(t))-\mathbf{R}(n^{-1}\mathbf{X}(t))|\le c_5(A(t)+I(t))^{-1}.$$

And –punch line– the first term in (38) equals zero, see (31). Therefore

$$\begin{aligned}&\mathbb{E}\left[\exp\left\{(n/\omega)^{1/2}\left[F_0(n^{-1}\mathbf{X}(t+1))-F_0(n^{-1}\mathbf{X}(t))\right]\right\}\Big|\circ\right] \\ \le& 1+c_4n^{-1}+O\big((n/\omega)^{1/2}n^{-1}(A(t)+I(t))^{-1}\big)\le 1+c_5n^{-1}.\end{aligned} \tag{39}$$

Combining (32) and (39) we get

$$\mathbb{E}[Q(t+1)|\circ]\le Q(t)(1+c_5n^{-1}),\quad \forall t<T. \tag{40}$$

(II) Define

$$\mathcal{Q}(t)=\begin{cases}Q(t)(1+c_5n^{-1})^{-t}, & t\le T,\\ Q(T)(1+c_5n^{-1})^{-T}, & t>T.\end{cases}$$

It follows from (40) that $\{\mathcal{Q}(t)\}_{t\ge 0}$ is a positive supermartingale. Given $z>0$, introduce $\mathcal{T}$ the first $t\le T$ such that $|F_0(n^{-1}\mathbf{X}(t))-F_0(n^{-1}\mathbf{X}(0))|>z$; set $\mathcal{T}=T+1$ if no such t exists. Obviously $\mathcal{T}$ is a stopping time. So by Optional Sampling Theorem (Durrett [17], Ch.4),

$$\mathbb{E}[\mathcal{Q}(\mathcal{T})]\le\mathcal{Q}(0)=Q(0)=1.$$

On the event $\{\mathcal{T} \le T\}$,

$$\mathcal{Q}(\mathcal{T}) \ge (1 + c_5 n^{-1})^{-\mathcal{T}} e^{z(n/\omega)^{1/2}} \ge e^{-c_5\mathcal{T}/n} e^{z(n/\omega)^{1/2}} \ge e^{-c_5} e^{z(n/\omega)^{1/2}}.$$

At the last step we used $\mathcal{T} \le n$; that's where the $O(n^{-1})$ remainder term came up so handy! So

$$\begin{aligned} &\mathbb{P}\left\{\max_{t\le T} |F_0(n^{-1}\mathbf{X}(t)) - F_0(n^{-1}\mathbf{X}(0))| > z\right\} \\ &\le \mathbb{P}\left\{\mathcal{Q}(\mathcal{T}) \ge e^{-c_5} e^{z(n/\omega)^{1/2}}\right\} \le \frac{\mathbb{E}[\mathcal{Q}(\mathcal{T})]}{e^{-c_5} e^{z(n/\omega)^{1/2}}} \\ &= e^{c_5} e^{-z(n/\omega)^{1/2}}. \end{aligned}$$

The same argument works for every F_j, $j < k$.

(III) Using the definition of F_js, we have therefore: with probability $1 - O(e^{-cz(n/\omega)^{1/2}})$,

$$\left| M(U(t))(n^{-1}\mathbf{I}(t)) - M(U(0))(n^{-1}\mathbf{I}(0)) \right| \le z.$$

Here $\mathbf{I}(t) = (I_0(t), \ldots, I_{k-1}(t))$, $U(t) = \ln(n^{-1}(A(t) + I(t)))^{-1/2}$, and $M(u)$ is the $k \times k$ matrix defined by the linear transformation (29). Since $M^{-1}(u)M(v) = M^{-1}(u - v)$, and the entries of $M^{-1}(u)$ are bounded for $u \in [0, \infty)$, we see that

$$\left| n^{-1}\mathbf{I}(t) - M^{-1}(V(t)) n^{-1}\mathbf{I}(0) \right| \le z,$$

with probability $1 - O(e^{-c_1 z(n/\omega)^{1/2}})$. Here

$$V(t) = \ln\left(\frac{n^{-1}A(0) + n^{-1}I(0)}{n^{-1}A(t) + n^{-1}I(t)}\right)^{1/2} = \ln(1 - 2t/nd)^{-1/2},$$

so that

$$M^{-1}(V(t)) n^{-1}\mathbf{I}(0) = \mathbf{i}(t/n),$$

with $\mathbf{i}(\tau)$ the solution of the linear differential equations (13) subject to the initial conditions $i_j(\tau)\big|_{\tau=0} = n^{-1}I_j(t)\big|_{t=0}$. Combining this with

$$n^{-1}A(t) + n^{-1}I(t) = d - 2t/n = a(t/n) + i(t/n),$$

we conclude that

$$\mathbb{P}\left\{\max_{t\le T} \left|n^{-1}\mathbf{X}(t) - \mathbf{x}(n^{-1}t)\right| \le z\right\} = 1 - O(e^{-c_2 z(n/\omega)^{1/2}}).$$

□

4 Phase transition window

4.1 Case $k < d-1$.

Corollary 5. *Let $p - p^* \ge \omega n^{-1/2}$, where $\omega = \omega(n) \to \infty$ sufficiently slowly, so that $\log n \gg \log \omega(n)$. Then, for some absolute constant $\alpha > 0$,*

$$\mathbb{P}\left\{I(T) \le n\omega^{-\frac{1}{d+2}}\right\} = 1 - O(\omega^{-\alpha}).$$

Consequently, for each $\varepsilon > 0$, w.h.p. the terminal number of inactive vertices is below εn.

Proof. Notice first that if $I(T) \ge n\omega^{-\frac{1}{d+2}}$ then, by the definition of T, we must have $A(T) = 0$ and $I(T) = nd - 2T$. Further, by Proposition 4 with $z = (\omega/n)^{1/2-\delta}$, $\delta = \frac{\log \omega}{3 \log n} \to 0$, for the event

$$B_n = \left\{\max_{j<k} \left|n^{-1} I_j(T) - i_j(n^{-1}T)\right| \le \left(\frac{\omega}{n}\right)^{1/2-\delta}\right\}$$

we have

$$\mathbb{P}(B_n) = 1 - O(\omega^{-\alpha}), \quad \alpha > 0.$$

On the event $B_n \cap \{I(T) \ge n\omega^{-\frac{1}{d+2}}\}$, by the definition of $i(\tau)$ and $S(y,p)$,

$$d\left(\frac{\omega}{n}\right)^{1/2-\delta} \ge n^{-1}I(T) - i(n^{-1}T) = dS(Y,p), \tag{41}$$
$$Y := \left(1 - \frac{2T}{nd}\right)^{1/2}.$$

Since $p - p^* \ge \omega n^{-1/2}$, using (22) we obtain

$$S(y(p), p) \ge c\omega n^{-1/2};$$

here $y(p)$ is the local minimum point of $S(y,p)$, ($y(p) \to y*$ as $p \downarrow p^*$). Observe that

$$\omega n^{-1/2} \gg \left(\frac{\omega}{n}\right)^{1/2-\delta},$$

as

$$\left(\frac{1}{2} + \delta\right) \log \omega - \delta \log n = \left(\frac{1}{2} + \delta - \delta \frac{\log n}{\log \omega}\right) \log \omega \ge \left(\frac{1}{2} - \frac{1}{3}\right) \log \omega \to \infty.$$

Then (41) and the properties of $S(y,p)$ imply that Y^2 is small, of order $(\omega/n)^{1/2-\delta}$, and so

$$I(T) = ndY^2 = O(n(\omega/n)^{1/2-\delta}) = O(n(\omega/n)^{1/3}) \ll n\omega^{-\frac{1}{d+2}}.$$

Thus we have proved that $I(T) \leq n\omega^{-\frac{1}{d+2}}$ on the event B_n, such that $\mathbb{P}(B_n) = 1 - O(\omega^{-\alpha})$. □

Lemma 6. *Let $m \leq m_n = o(n)$. The random d-regular multigraph on n vertices has no submultigraph on m vertices with $3m/2$ or more edges with probability $1 - O(n^{-1})$.*

Proof. We use the equivalent configuration model. There are n disjoint sets $S_1, \dots, S_n$, each of cardinality d, and we consider the random pairing chosen uniformly at random among all $(nd-1)!!$ pairings on $\cup_i S_i$. Let us compute $\mathbb{E}(m,r)$ the expected number of unordered collections $S_{i_1}, \dots, S_{i_m}$ such that there are exactly r pairs of points in $\cup S_{i_j}$. There are $\binom{n}{m}$ collections $S_{i_1}, \dots, S_{i_m}$. Given such a collection, we pick $2r$ points from their union in $\binom{md}{2r}$ ways, and pair them in $(2r-1)!!$ ways. The remaining $md-2r$ points of these sets are matched with some $md-2r$ points from the remaining $n-m$ sets S_i, and this step can be done in

$$\binom{(n-m)d}{md-2r}(md-2r)! = \frac{[(n-m)d]!}{[(n-2m)d+2r]!}$$

ways. Finally we pair the remaining $(n-2m)d+2r$ points in $[(n-2m)d + 2r - 1]!!$ ways. Multiplying all the numbers, dividing by the total number of pairings, $(nd-1)!!$, and using $(2m-1)!! = (2m)!/2^m m!$, we obtain

$$\mathbb{E}(m,r) = \binom{n}{m} \frac{(md)!}{2^{2r-md}\, r!(md-2r)!} \frac{[(n-m)d]!}{(nd/2 - md + r)!} \frac{(nd/2)!}{(nd)!}.$$

Notice that $r \leq md/2 \leq m_n d/2 = o(n)$. For this range of m, r, the above expression yields

$$\mathbb{E}(m,r) \leq (1 + O(m_n/n))^m E^*(m,r); \quad E^*(m,r) := n^{m-r} \frac{(md)!}{m!r!(md-2r)!}.$$

Observe that, for r, m in question,

$$\frac{E^*(m,r+1)}{E^*(m,r)} = O(n^{-1}m^2/r) = O(m_n/n) = o(1),$$

so

$$\sum_{r\geq 3m/2}\mathbb{E}(m,r)\leq (1+O(m_n/n))^{3m/2}E^*(m,3m/2).$$

Furthermore

$$\frac{(md)!}{m!(3m/2)!(md-3m)!}=\frac{(md)!}{(md-m/2)!}\binom{md-m/2}{m,3m/2,md-3m}\leq (md)^{m/2}3^{md},$$

so, as $m_n=o(n)$,

$$\sum_{1\leq m\leq n}(1+O(m_n/m))^{3m/2}E^*(m,3m/2)$$
$$\leq O(n^{-1})+\sum_{2\leq m\leq m_n}\left(\frac{(md)^{1/2}4^d}{n^{1/2}}\right)^m=O(n^{-1}).$$

Hence the expected number of submultigraphs with the number of vertices $m\leq m_n$ falling below the number of edges by a factor 2/3 or less is of order n^{-1}. So, with probability $1-O(n^{-1})$, no such submultigraph is present. □

Corollary 7. *For $p-p^*\geq \omega n^{-1/2}$, with probability $1-O(\omega^{-a})$ no inactive vertex remains at the end of the process.*

Proof. By Corollary 5, with probability $1-O(\omega^{-a})$, $I(T)$ the number of inactive points at the stopping time T, is $n\omega^{-1/(d+2)}$ at most, and then so is the size of $\mathcal{I}_f$, the *final* set of inactive vertices. Such a set, if present, induces a submultigraph of minimum degree at least $d-k+1\geq 3$, thus having at least $1.5|\mathcal{I}_f|$ edges. For $|\mathcal{I}_f|\leq n\omega^{-1/(d+2)}=o(n)$, by Lemma 6, the probability of the last event is $O(n^{-1})$. Hence the probability of incomplete activation is of order $\omega^{-a}+n^{-1}\to 0$. □

Corollary 8. *Let $p-p^*\leq -\omega n^{-1/2}$, where $\omega=\omega(n)\to\infty$ sufficiently slowly, so that $\log n\gg\log\omega(n)$. Let $\hat{y}(p)$ denote the larger positive root of $S(y,p)=0$, ($\hat{y}(p)\to y^*$, $p\to p^*$), and $\hat{\tau}(p)$ denote the smaller root of $a(\tau)=0$, $\hat{\tau}(p)=(d/2)(1-\hat{y}(p)^2)$. Then, for each $\gamma>1/2$, with probability*

$1 - O(\omega^{-\alpha})$, $(\alpha = \alpha(\gamma) > 0)$,

$$\begin{aligned} A(T) =& 0, \\ \sum_{j<k} I_j(T) =& n\mathbb{P}(Bin(d, 1-\hat{y}(p)) < k) + O\big(n^{1/2}\omega^{\gamma}(p^* - p)^{-1/2}\big), \\ T =& \frac{nd}{2}(1-\hat{y}(p)^2) + O\big(n^{1/2}\omega^{\gamma}(p^* - p)^{-1/2}\big) \\ =& n\hat{\tau}(p) + O\big(n^{1/2}\omega^{\gamma}(p^* - p)^{-1/2}\big). \end{aligned}$$

Note. Thus w.h.p. the stopping time T is the termination time, $n^{-1}T$ is asymptotic to the smaller root $\hat{\tau}(p)$ of $a(\tau) = 0$, and $n^{-1}I_f$ is asymptotic to $\mathbb{P}(\mathrm{Bin}(d, 1-\hat{y}(p)) < k)$.

Proof. The argument runs parallel to the proof of Corollary 5. First of all, since now $p - p^* \le -\omega n^{-1/2}$,

$$S(y(p), p) = \min\{S(y, p) : \, y \in [y(p), 1]\} \le -c\omega n^{-1/2}.$$

Introduce $\tau(p) = (d/2)(1 - y(p)^2)$. By the definitions, $a(\tau(p)) = dS(y(p), p)$. Set $T(p) = [n\tau(p)]$. Intuitively it is unlikely that $T \ge T(p)$ since otherwise $n^{-1}A(T(p))$ would be close to $a(\tau(p))$, which is negative! Let us make it rigorous. As in the proof of Corollary 5,

$$\begin{aligned} \mathbb{P}(B_n) =& 1 - O(\omega^{-\alpha}), \\ B_n :=& \left\{ \max_{t \le T,\, j<k} \left| n^{-1}I_j(t) - i_j(n^{-1}t) \right| \le \left(\frac{\omega}{n}\right)^{1/2-\delta} \right\}, \quad \delta = \beta \frac{\log \omega}{\log n}. \end{aligned}$$

On the event $B_n \cap \{T \ge T(p)\}$, we have

$$\begin{aligned} n^{-1}A(T(p)) \le& n^{-1}(nd - 2T(p)) - i(n^{-1}T(p)) + c_1 \left(\frac{\omega}{n}\right)^{1/2-\delta} \\ =& dS(y(p), p) + c_1 \left(\frac{\omega}{n}\right)^{1/2-\delta} \\ \le& -c_2\omega n^{-1/2} + c_1 \left(\frac{\omega}{n}\right)^{1/2-\delta}, \end{aligned}$$

and the last expression is negative for n large, by the definition of δ. Thus $B_n \cap \{T \ge T(p)\}$ is empty, and $B_n \subseteq \{T \le T(p)\}$.

Further, on the event $\{T < T(p)\}$,

$$\begin{aligned} A(T) + I(T) =& nd - 2T \geq nd - 2T(p) \\ =& nd\left(1 - \frac{2\tau(p)}{d}\right) + O(1) = ndy(p)^2 + O(1), \end{aligned}$$

which is of order n exactly. Likewise

$$\begin{aligned} n^{-1}I(T) \geq& i(n^{-1}T) - c_3\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ \geq& (1-p)\left(1 - \frac{2T}{nd}\right)^{d/2} - c_3\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ \geq& (1-p)\left(1 - \frac{2T(p)}{nd}\right)^{d/2} - c_3\left(\frac{\omega}{n}\right)^{1/2-\delta}, \end{aligned}$$

so that $I(T)$ is of an exact order n as well. Therefore $A(T) = 0$ on the event B_n.

Assume that B_n happened. We already know that $T < T(p)$. Having proved that $A(T) = 0$ allows us to show that in fact $T \sim n\hat{\tau}(p)$ and I_f, the size of the final inactive set $\mathcal{I}_f$ is close to $n\mathbb{P}(\mathrm{Bin}(d, 1-\hat{y}(p)) < k)$.
First,

$$\begin{aligned} 0 = n^{-1}A(T) =& d - \frac{2T}{n} - i(n^{-1}T) + O\left[\left(\frac{\omega}{n}\right)^{1/2-\delta}\right] \\ =& dS(Y,p) + O\left[\left(\frac{\omega}{n}\right)^{1/2-\delta}\right]; \quad Y = \left(1 - \frac{2T}{nd}\right)^{1/2}. \end{aligned} \tag{42}$$

Here

$$Y \geq \left(1 - \frac{2T(p)}{nd}\right)^{1/2} = \left(1 - \frac{2\tau(p)}{d}\right)^{1/2} = y(p),$$

where $y(p)$ is the minimum point of $S(y,p)$. (42) means that $S(Y,p)$ is of order $(\omega/n)^{1/2-\delta}$, so that Y must be close to $\hat{y}(p)$, the root of $S(y,p) = 0$ to the right of $y(p)$. Indeed, since $S_y(y(p), p) = 0$,

$$S(y,p) \leq S(y(p),p) + c_1(y - y(p))^2, \quad c' = 0.5\max\{|S_{yy}(y,\pi) : \, y, \pi \in [0,1]\}.$$

So, using this inequality for $y = \hat{y}(p)$ and $y = Y$ together with $S(\hat{y}(p), p) = 0$, (42), and (22),

$$Y - y(p),\, \hat{y}(p) - y(p) \geq c''(p^* - p)^{1/2}. \tag{43}$$

Now notice that by (21), for $y \geq y(p) + c''(p^* - p)^{1/2}$,

$$S_y(y, p) \geq S_y(y(p), p) + c_1(y - y(p)) \geq c_1 c''(p^* - p)^{1/2}. \tag{44}$$

Combining (43), (44) with $S(\hat{y}(p), p) = 0$, (42), we obtain

$$|Y - \hat{y}(p)| \leq c_2 \frac{(\omega/n)^{1/2-\delta}}{(p^* - p)^{1/2}}.$$

Therefore

$$\begin{aligned} T =& \frac{nd}{2}(1 - Y^2) = \frac{nd}{2}(1 - \hat{y}(p)^2) + O\big(n^{1/2+\delta}\omega^{1/2-\delta}(p^* - p)^{-1/2}\big) \\ =& n\hat{\tau}(p) + O\big(n^{1/2}\omega^{1/2+\beta}(p^* - p)^{-1/2}\big). \end{aligned}$$

Consequently, by the definition of the event B_n,

$$\begin{aligned} I_f =& \sum_{j<k} I_j(T) \\ =& n \sum_{j<k} \big[(1 - p)i_j(n^{-1}T) + O((\omega/n)^{1/2-\delta})\big] \\ =& n(1 - p) \sum_{j<k} i_j(\hat{\tau}(p)) + O\big(n^{1/2}\omega^{1/2+\beta}(p^* - p)^{-1/2}\big) \\ =& n(1 - p)\mathbb{P}(\mathrm{Bin}(d, 1 - \hat{y}(p)) < k) + O\big(n^{1/2}\omega^{1/2+\beta}(p^* - p)^{-1/2}\big). \quad \square \end{aligned}$$

Proof of Theorem 1. Corollaries 7 and 8 together constitute Theorem 1. □

4.2 Case $k = d - 1$.

Corollary 9. *Let $p \geq p^*$. If $\omega \to \infty$ and $\log \omega \ll \log n$ then, for some absolute constant $\alpha > 0$,*

$$\mathbb{P}\left\{ I(T) \leq n\omega^{-\frac{1}{d+2}} \right\} = 1 - O(\omega^{-\alpha}).$$

Proof. For $p \in [p^*, 1)$ and $y \in (0, 1)$,

$$\begin{aligned} S(y, p) =& y^2 \left[1 - (1 - p)(d - 1) + (1 - p)\frac{(1 - y)^{d-1} - (1 - (d - 1)y)}{y} \right] \\ =& y^2 \left[\frac{p - p^*}{1 - p^*} + (1 - p)\frac{(1 - y)^{d-1} - (1 - (d - 1)y)}{y} \right] \qquad (45) \\ \geq& cy^3, \quad c = (1 - p)2^{3-d} \cdot \binom{d-1}{2}. \end{aligned}$$

So (see the proof of Corollary 5) on the event $B_n \cap \left\{I(T) > n\omega^{-\frac{1}{d+2}}\right\}$, we have

$$d\left(\frac{\omega}{n}\right)^{1/2-\delta} \geq n^{-1}I(T) - i(n^{-1}T) \geq dS(Y,p) \geq cdY^3,$$

$Y = (1-2T/nd)^{1/2}$. Therefore

$$\begin{aligned} n^{-1}I(T) \leq i(n^{-1}T) + d\left(\frac{\omega}{n}\right)^{1/2-\delta} &= dY(1-p)[1-(1-Y)^{d-1}] + d\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ &\leq d^2(1-p)Y^2 + d\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ &\leq c_1\left(\frac{\omega}{n}\right)^{1/3-2\delta/3}, \end{aligned}$$

so that

$$I(T) \leq c_1 n\left(\frac{\omega}{n}\right)^{1/3-2\delta/3} \leq c_1 n^{2/3}\omega^{1/3}.$$

It follows then that $B_n \cap \left\{I(T) > n\omega^{-\frac{1}{d+2}}\right\} = \emptyset$, so that

$$\mathbb{P}\left\{I(T) \leq n\omega^{-\frac{1}{d+2}}\right\} \geq \mathbb{P}(B_n) = 1 - O(\omega^{-\alpha}).$$

□

Since $I(t)$ decreases with t we obtain that w.h.p. the final set $\mathcal{I}_f$ of inactive vertices has cardinality $I_f = O(n\omega^{-\frac{1}{d+2}})$, for $\omega \to \infty$ sufficiently slowly. Since $k = d-1$, the submultigraph induced by $\mathcal{I}_f$ has minimum degree at least 2.

Lemma 10. *Let $p - p^* > 0$, and let $\lambda = \lambda(n) \to \infty$ however slowly. Let S_n denote the total size of all induced submultigraphs of minimum degree 2 or more on initially inactive vertices, each having at most $\ell_n = n\lambda^{-1}(p-p^*)^2$ vertices. Then*

$$\mathbb{E}[S_n] = O((p-p^*)^{-3/2}).$$

Consequently,

$$S_n = O_P((p-p^*)^{-3/2}).$$

Proof. Again we work with the configuration model. Given $\ell \leq \ell_n$, let $N_{n,\ell}$ denote the total number of induced submultigraphs of minimum degree 2 or more, containing ℓ initially inactive vertices. Let us compute $\mathbb{E}[N_{n,\ell}]$. A collection of ℓ initially inactive vertices, i. e. the sets $S_{i_1}, \dots, S_{i_\ell}$, can be chosen in $\binom{n(1-p)}{\ell}$ ways, as there are exactly $n(1-p)$ of initially inactive vertices, i. e. sets S_i. Then we select from each S_{i_j} some $\delta_j \in [2, d]$ points such that $\Delta = \sum_j \delta_j$ is even; this can be done in

$$\prod_{j=1}^{\ell} \binom{d}{\delta_j}$$

ways. Lastly, we match the selected Δ points among themselves, in $(\Delta - 1)!!$ ways, and then match the remaining $(nd - \Delta)$ points among themselves, in $(nd - \Delta - 1)!!$ ways. Consequently,

$$\mathbb{E}[N_{n,\ell}] = \binom{n(1-p)}{\ell} \sum_{2 \le \delta_1, \dots, \delta_\ell \le d} \prod_{j=1}^{\ell} \binom{d}{\delta_j} \frac{(\Delta - 1)!!(nd - \Delta - 1)!!}{(nd - 1)!!}.$$

Therefore

$$\begin{aligned} \mathbb{E}[N_{n\ell}] \le & \frac{\binom{n(1-p)}{\ell}}{(nd-1)!!} \sum_{\Delta = 2\ell}^{d\ell} Q(\Delta), \\ Q(\Delta) := & (\Delta - 1)!!(nd - \Delta - 1)!! \sum_{\substack{2 \le \delta_j \le d \\ \delta_1 + \dots + \delta_\ell = \Delta}} \prod_{j=1}^{\ell} \binom{d}{\delta_j}. \end{aligned}$$

We need a sharp bound for the last sum. Generating functions and Chernoff's method to the rescue! Denoting $f(x) = \sum_{j=2}^{d} \binom{d}{j} x^j$,

$$\sum_{\substack{2 \le \delta_j \le d \\ \delta_1 + \dots + \delta_\ell = \Delta}} \prod_{j=1}^{\ell} \binom{d}{\delta_j} = [x^\Delta] f(x)^\ell \le \frac{f(x)^\ell}{x^\Delta}, \qquad \forall\, x > 0.$$

Therefore

$$\begin{aligned} \mathbb{E}[N_{n,\ell}] \le & \frac{\binom{n(1-p)}{\ell}}{(nd-1)!!} \sum_{\Delta = 2\ell}^{d\ell} \tilde{Q}(\Delta), \\ \tilde{Q}(\Delta) := & (\Delta - 1)!!(nd - \Delta - 1)!! \frac{f(x)^\ell}{x^\Delta}, \quad \forall\, x > 0. \end{aligned}$$

Now, using $\Delta \le d\ell \le nd\lambda^{-1}(p - p^*)^2$,

$$\frac{\tilde{Q}(\Delta + 2)}{\tilde{Q}(\Delta)} = \frac{\Delta + 1}{(nd - \Delta - 1)x^2} \le \frac{1}{2},$$

if

$$x = c \cdot \frac{p - p^*}{\lambda^{1/2}},$$

and c is chosen sufficiently large. For this x, we get then

$$\mathbb{E}[N_{n,\ell}] \leq E_{n,\ell} := 2\frac{\binom{n(1-p)}{\ell}}{(nd-1)!!}\tilde{Q}(2\ell).$$

Notice that

$$\frac{f(x)^\ell}{x^{2\ell}} = \binom{d}{2}^\ell (1+O(x))^\ell,$$

and

$$\begin{aligned}\frac{(nd-2\ell-1)!!}{(nd-1)!!} =&(1+O(\ell/n))^\ell (nd)^{-\ell} = (1+O(x^2))^\ell (nd)^{-\ell},\\ \binom{n(1-p)}{\ell} =&(1+O(\ell/n)))^\ell \frac{n^\ell(1-p)^\ell}{\ell!} = (1+O(x^2))^\ell \frac{n^\ell(1-p)^\ell}{\ell!}. \quad (46)\end{aligned}$$

Putting things together,

$$\begin{aligned}E_{n,\ell} =&c_1(1+O(x))^\ell \frac{(2\ell-1)!!}{2^\ell \ell!}(d-1)^\ell(1-p)^\ell\\ =&c_1(1+O(x))^\ell \frac{\binom{2\ell}{\ell}}{2^{2\ell}}\left(\frac{1-p}{1-p^*}\right)^\ell.\end{aligned}$$

The fraction $2^{-2\ell}\binom{2\ell}{\ell}$ is of order $\ell^{-1/2}$. Since

$$\frac{1-p}{1-p^*} = 1 - \frac{p-p^*}{1-p^*},$$

and $x = o(p - p^*)$, we see that $\sum_{\ell \leq \ell_n} \ell E_{n,\ell}$ is of order

$$\begin{aligned}\sum_{\ell\geq 1} \ell^{1/2}\left((1+O(x))\frac{1-p}{1-p^*}\right)^\ell =&O\left[\left(1-\frac{1-p}{1-p^*}\right)^{-3/2}\right]\\ =&O((p-p^*)^{-3/2}).\end{aligned}$$

□

Corollary 9 and Lemma 10 can be combined as follows.

Corollary 11. *Suppose* $p - p^* \geq n^{-\varepsilon}$, *where* $\varepsilon = \varepsilon(n) \downarrow 0$, *and* $\varepsilon \log n \to \infty$. *Then w.h.p. the terminal inactive set contains at most* $O((p-p^*)^{-3/2})$ *vertices.*

Proof. By Corollary 9, with probability $1 - O(\omega^{-a})$, $I(T)$ the number of inactive points at the stopping time T, is $n\omega^{-\frac{1}{d+2}}$ at most. Then so is I_f, the size of the final set of inactive vertices. Such a set, if present, induces a submultigraph of minimum degree at least $d - k + 1 \geq 2$. Suppose that ω and λ are such that

$$n\omega^{-\frac{1}{d+2}} \leq n\lambda^{-1}(p-p^*)^2. \tag{47}$$

Then by Lemma 10 w.h.p. $I_f \leq O((p-p^*)^{-3/2})$. If $p - p^* \geq n^{-\varepsilon}$ then (47) satisfied for $\lambda = n^{\varepsilon}$ and $\omega = n^{3\varepsilon(d+2)}$, and the conditions of Corollary 9 and Lemma 10 are satisfied as well. □

Corollary 12. *Let* $p - p^* \leq -\lambda n^{-\sigma}$, $\sigma = \frac{1}{2(d+5)}$, *with* $\lambda = \lambda(n) \to \infty$, $\log \lambda = o(\log n)$. *Then the statement of Corollary 8 holds, with the remainder term* $O(n^{1-3\sigma}\lambda^{6\sigma}(p^*-p)^{-1})$.

Proof. As in the proof of Corollary 8, we introduce

$$\tau(p) = \frac{d}{2}(1 - y(p))^2, \quad T(p) = [n\tau(p)],$$

where $y(p)$ is the minimum point of $S(y,p)$. According to (24) $y(p)$ is of order $p^* - p$, and $-S(y(p),p)$ is of order $(p^*-p)^3$ if $p \uparrow p^*$. Let B_n be the event used in the proof of Corollary 5, with $\omega \to \infty$ ($\omega = o(n)$) to be determined later and $\delta = \frac{\log \lambda}{\log n}$. By Proposition 4, $\mathbb{P}(B_n) = 1 - O(\lambda^{-\alpha})$, $\alpha > 0$. On the event $B_n \cap \{T \geq T(p)\}$, under the assumption

$$(p^* - p)^3 \gg \left(\frac{\omega}{n}\right)^{1/2-\delta} \tag{48}$$

we have

$$\begin{aligned} n^{-1}A(T(p)) \leq & dS(y(p),p) + c_1\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ \leq & -c_2(p^*-p)^3 + c_1\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ < & 0. \end{aligned}$$

Therefore $B_n \subseteq \{T < T(p)\}$. Further, on $\{T < T(p)\}$,

$$\begin{aligned} A(T) + I(T) =& nd - 2T \geq nd - 2T(p) \\ =& nd\left(1 - \frac{2\tau(p)}{d}\right) + O(1) \\ =& ndy(p)^2 + O(1), \end{aligned}$$

so that $A(T)+I(T)$ is of the exact order $n(p^*-p)^2$. Next, for $y = (1-2\tau/d)$,

$$i(\tau) = dy(1-p)\left[1-(1-y)^{d-1}\right] \geq cy^2.$$

So, since $I(T) \geq I(T(p))$,

$$\begin{aligned} n^{-1}I(T) \geq& i(n^{-1}T(p)) - c_3\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ \geq& c\left(1 - \frac{2T(p)}{nd}\right) - c_3\left(\frac{\omega}{n}\right)^{1/2-\delta} \\ \geq& c'(p^*-p)^2 - c_3\left(\frac{\omega}{n}\right)^{1/2-\delta}. \end{aligned}$$

which means that $I(T)$ is of the exact order $n(p^*-p)^2$ too, if

$$(p^*-p)^2 \gg \left(\frac{\omega}{n}\right)^{1/2-\delta}. \tag{49}$$

But notice that if

$$(p^*-p)^2 \gg \omega^{-\frac{1}{d+2}}, \tag{50}$$

then

$$I(T) \gg n\omega^{-\frac{1}{d+2}}.$$

Thus on the event $\{T < T(p)\}$, $I(T)$ far exceeds $n\omega^{-\frac{1}{d+2}}$. By the definition of the stopping time T we must have $A(T) = 0$ on $\{T < T(p)\}$.

To comply with the restrictions (48), (49), (50) we have to choose ω such that

$$p^* - p \gg \max\{(\omega^{-\frac{1}{d+2}})^{1/2}, \left(\frac{\omega}{n}\right)^{(1/2-\delta)/2}, \left(\frac{\omega}{n}\right)^{(1/2-\delta)/3}\}.$$

A closer look at this restriction shows that our best choice is to set $\omega = n^{\frac{d+2}{d+5}}$, which also explains the condition $p^* - p \geq \lambda n^{-\frac{1}{2(d+5)}}$, with $\lambda = \lambda(n) \to \infty$, $\log\lambda = o(\log n)$.

Suppose the event B_n has happened. We have proved that then necessarily $T < T(p)$ and $A(T) = 0$. Then, see (42),

$$S(Y,p) = O\left[\left(\frac{\omega}{n}\right)^{1/2-\delta}\right], \quad Y = \left(1 - \frac{2T}{nd}\right)^{1/2} \geq y(p),$$

so that, by (28),

$$c_4(-(p^*-p)^3 + (p^*-p)(Y-y(p))^2 + (Y-y(p))^3) \geq O\left[\left(\frac{\omega}{n}\right)^{1/2-\delta}\right].$$

This relation and (48) imply that

$$Y - y(p) \geq c_5(p^* - p).$$

Recall, (23), that

$$\hat{y}(p) - y(p) \geq c_6(p^* - p).$$

Using (27) we see that

$$\min\{S_y(y,p) : p \in [Y, \hat{y}(p)]\} \geq c^*(p^* - p)^2.$$

Combining this estimate with $S(\hat{y}(p), p) = 0$ and $S(Y,p) = O((\omega/n)^{1/2-\delta})$ we conclude that

$$Y = \hat{y}(p) + O\left[\left(\frac{\omega}{n}\right)^{1/2-\delta}(p^*-p)^{-2}\right].$$

Consequently

$$\begin{aligned} T =& \frac{nd}{2}(1 - Y^2) \\ =& \frac{nd}{2}(1 - \hat{y}^2(p)) + O\left[n\left(\frac{\omega}{n}\right)^{1/2-\delta}(p^*-p)^{-1}\right] \\ =& n\hat{\tau}(p) + o(n^{1-\frac{3}{2(d+5)}+o(1)}(p^*-p)^{-1}). \end{aligned}$$

Also

$$\begin{aligned} I_f =& \sum_{j<k} I_j(T) = n\sum_{j<k}\left[i_j(T/n) + O((\omega/n)^{1/2-\delta})\right] \\ =& n\sum_{j<k}\left[(1-p)i_j(\hat{\tau}(p)) + O(n^{-\frac{3}{2(d+5)}}\lambda^{\frac{3}{d+5}}(p^*-p)^{-1})\right] \\ =& n(1-p)\mathbb{P}(\mathrm{Bin}(d, 1-\hat{y}(p)) < k) + O(n^{1-3\sigma}\lambda^{6\sigma}(p^*-p)^{-1}). \end{aligned}$$

Here

$$
\begin{aligned}
n\mathbb{P}(\mathrm{Bin}(d, 1-\hat{y}(p) < k) \sim & n\binom{d}{2}\hat{y}^2(p) \geq cn(p^* - p)^2 \\
\gg & n^{1-3\sigma}\lambda^{6\sigma}(p^* - p)^{-1},
\end{aligned}
$$

as $p^* - p \geq \lambda n^{-\sigma}$. □

Proof of Theorem 2. Corollaries 11 and 12 constitute Theorem 2 (i) and (ii). □

References

[1] M. Aizenman and J. L. Lebowitz, Metastability effects in bootstrap percolation. *J. Phys. A* **21** (1988), 3801–3813.

[2] D. Aldous and B. Pittel, On a random graph with immigrating vertices: emergence of the giant component, *Random Structures and Algorithms* **17** (2000), 79–102.

[3] J. Aronson, A. Frieze, B. G. Pittel, Maximum matchings in sparse random graphs: Karp-Sipser revisited, *Random Structures and Algorithms* **12** (1998), 111–177.

[4] J. Balogh and G. Pete, Random disease on the square grid. *Random Structures and Algorithms* **13** (1998), 409–422.

[5] J. Balogh and B. Bollobás, Sharp thresholds in bootstrap percolation, *Physics A* **326** (2003), 305–312.

[6] J. Balogh and B. Bollobás, Bootstrap percolation on the hypercube, *Probability Theory and Related Field* to appear.

[7] J. Balogh, Y. Peres and G. Pete, Bootstrap percolation on infinite trees and non-amenable groups, *Combinatorics, Probability and Computing* to appear.

[8] E. Bender and R. Canfield, The asymptotic number of labelled graphs with given degree sequences, *J. Comb. Theory, Ser. A* **24** (1978), 296–307.

[9] E.R. Berlekamp, J.H. Conway and R.K. Guy, *Winning Ways for Your Mathematical Plays*, vol. 2, Academic Press. London, 1982.

[10] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Preprint Series, Mathematisk Institut, Aarhus Universitet.*

[11] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal of Combinatorics* **1** (1980), 311–316.

[12] B. Bollobás, The isoperimetric number of random regular graphs, *European J. Combin.* **9** (1988), 241–244.

[13] B. Bollobás, Private communication (2001).

[14] A. W. Burks (ed.), *Essays on Cellular Automata*, University of Illinois Press, 1970.

[15] R. Cerf and Manzo, The threshold regime of finite volume bootstrap percolation, *Stochastic Process. Appl.* **101** (2002), 69–82.

[16] J. Chalupa, G. R. Reich and P. L. Leath, Bootstrap percolation on a Bethe Lattice, *J. Phys. C* **12** (1979), L31–L35.

[17] R. Durrett, Probability: theory and examples. Second edition. Duxbury Press, Belmont, CA, 1996. xiii+503 pp.

[18] A. C. D. van Enter, Proof of Straley's argument for bootstrap percolation, *J. Statist. Phys.* **48** (1987), 943–945.

[19] E. Friedgut and G. Kalai, Every monotone graph property has a sharp threshold, *Proc. Amer. Math. Soc.* **124** (1996), 2993–3002.

[20] M. Gardner, *Wheels, Life and Other Mathematical Amusements*, W.H. Freeman and Co., 1983.

[21] G. R. Grimmett, Percolation. Second edition. Springer 1999.

[22] A. E. Holroyd, Sharp metastability threshold for two-dimensional bootstrap percolation, *Probability Theory and Related Fields* **125** (2003), 195–224.

[23] R. M. Karp and M. Sipser, Maximum matchings in sparse random graphs. *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computing* (1981), 364–375.

[24] T. M. Liggett, Interacting particle systems. Reprint of the 1985 original. Classics in Mathematics. Springer-Verlag, Berlin, 2005. xvi+496 pp.

[25] W. S. McCulloch and W. Pitt, A logical calculus of ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics* **5** (1943), 115–133.

[26] B. G. Pittel, J. Spencer, and N. Wormald, Sudden emergence of a giant k-core in a random graph, *J. Combin. Theory Ser. B* **67** (1996), 111–151.

[27] R. H. Schonmann, On the behavior of some cellular automata related to bootstrap percolation, *Ann. Probab.* **20** (1992), 174–193.

[28] S. Ulam, Random processes and transformations, in *Proceedings of the International Congress of Mathematicians*, Cambridge, Mass., 1950, Amer. Math. Soc., Providence, R.I., vol. 2, 1952, pp. 264–275.

[29] N. Wormald, The asymptotic connectivity of labelled regular graphs. *J. Comb. Theory, Ser. B* **31** (1981), 156–167.

[30] N. Wormald, The asymptotic distribution of short cycles in random regular graphs. *J. Comb. Theory, Ser. B* **31** 168–182.