

# Components, large and small, are as they should be II: supercritical percolation on regular graphs of constant degree

Sahar Diskin \*

Michael Krivelevich \*

## Abstract

Let  $d \geq 3$  be a fixed integer. Let  $y := y(p)$  be the probability that the root of an infinite  $d$ -regular tree belongs to an infinite cluster after  $p$ -bond-percolation. We show that for all constants  $b, \alpha > 0$  and  $1 < \lambda < d - 1$ , there exist constants  $c, C > 0$  such that the following holds. Let  $G$  be a  $d$ -regular graph on  $n$  vertices, satisfying that for every  $U \subseteq V(G)$  with  $|U| \leq \frac{n}{2}$ ,  $e(U, U^c) \geq b|U|$  and for every  $U \subseteq V(G)$  with  $|U| \leq \log^C n$ ,  $e(U) \leq (1 + c)|U|$ . Let  $p = \frac{\lambda}{d-1}$ . Then, with probability tending to one as  $n$  tends to infinity, the largest component  $L_1$  in the random subgraph  $G_p$  of  $G$  satisfies  $\left|1 - \frac{|L_1|}{yn}\right| \leq \alpha$ , and all the other components in  $G_p$  are of order  $O\left(\frac{\lambda \log n}{(\lambda-1)^2}\right)$ . This generalises (and improves upon) results for random  $d$ -regular graphs.

## 1 Introduction

Given a *host graph*  $G = (V, E)$  and a probability  $p \in [0, 1]$ , the random subgraph  $G_p$  is obtained by performing independent  $p$ -bond percolation on  $G$ , that is, we form  $G_p$  by retaining each edge in  $E$  independently with probability  $p$ . Perhaps the most well-studied example is the binomial random graph  $G(n, p)$ , which is equivalent to percolation with probability  $p$  on the complete graph  $K_n$ . For more background on random graphs and on percolation, we refer the interested reader to [8, 11, 20, 23, 25, 27].

A classical result of Erdős and Rényi from 1960 [17] states that  $G(n, p)$  undergoes a fundamental phase transition, with respect to its connected component structure, when the expected average degree is around one (that is,  $p \cdot (n - 1) \approx 1$ ): in the *subcritical regime*, when the expected average degree is less than one, all the components are typically of order  $O(\log n)$ , whereas in the *supercritical regime*, when the expected average degree is larger than one, there likely emerges a unique giant component, taking linear (in  $n$ ) fraction of the vertices, and all other components are typically of order  $O(\log n)$ . In fact, it is known that in the supercritical regime in  $G(n, p)$ , the asymptotic relative size of the giant component is dictated by the survival probability of a Galton-Watson process with offspring distribution  $\text{Bin}(n - 1, p)$ .

A quantitatively similar behaviour has been observed in percolation on several (families of)  $d$ -regular graphs, when  $d = \omega_n(1)$ : pseudo-random  $(n, d, \lambda)$ -graphs with  $\lambda = o(d)$  [21], the  $d$ -dimensional hypercube  $Q^d$  [1, 10], and other  $d$ -regular ‘high-dimensional’ graphs [13, 14, 29]. In a companion paper [15], the authors consider percolation on  $d$ -regular graphs  $G$  with  $d = \omega(1)$ , and provide sufficient and essentially tight conditions guaranteeing a phase transition in  $G_p$  quantitatively similar to that of  $G(n, p)$  when the expected average degree is around one. Roughly, it suffices to require  $G$  to satisfy a (very) mild ‘global’ edge expansion, and to have a fairly good control on the expansion of small sets in  $G$ . For the case of growing degree, this serves as a unified proof for the aforementioned ‘concrete’ host graphs.

In this paper, we consider  $d$ -regular graphs  $G$  on  $n$  vertices, where  $d \geq 3$  is a fixed integer and our asymptotics are in  $n$ .<sup>1</sup> Our goal is to obtain comparable (to the growing degree case) sufficient conditions on  $G$ , such that  $G_p$  exhibits a phase transition with respect to its component structure ‘similar’ to that of  $G(n, p)$  — that is, the typical emergence (when the expected average degree is above one) of a unique giant component taking linear number of the vertices, whereas all other components are typically much smaller, of order logarithmic in the number of vertices.

---

\*School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel.

Emails: sahardiskin@mail.tau.ac.il, krivelev@tauex.tau.ac.il.

<sup>1</sup>When  $d = 2$ , the graph  $G$  is just a collection of cycles.

Note that for a fixed  $d \geq 3$ , given a  $d$ -regular graph  $G$  and  $p = \frac{\lambda}{d-1}$  with  $\lambda < 1$ , it is known that **whp** all components of  $G_p$  are of order  $O(\log n)$  (see [31]). We thus focus our attention on the supercritical regime, that is, when  $p = \frac{\lambda}{d-1}$  with  $\lambda > 1$ .

Before stating our main result, let us give some intuition and discuss previous related results. For a fixed integer  $d \geq 3$  and  $p \in (0, 1)$ , let  $q := q(p)$  be the unique solution in  $(0, 1)$  of the equation

$$q = (1 - p + pq)^{d-1}. \quad (1)$$

This is the extinction probability of a Galton-Watson tree with offspring distribution  $\text{Bin}(d-1, p)$  (see, for example, [6, 16]). Consider then the probability  $y := y(p)$  that the root of an infinite  $d$ -regular tree belongs to an infinite cluster after  $p$ -bond-percolation. Then,  $y$  is given by

$$y = \sum_{i=1}^d \binom{d}{i} p^i (1-p)^{d-i} (1-q^i) = 1 - (1-p+pq)^d = 1 - q(1-p) - pq^2, \quad (2)$$

where  $q = q(p)$  is defined according to (1). As we will shortly see, and as might be anticipated,  $y$  will serve us both in the subsequent discussion and in the statement of our main result. We note that when  $d$  tends to infinity and  $p = \frac{1+\epsilon}{d-1}$  for small constant  $\epsilon > 0$ , then  $y$  is asymptotically equal to the unique solution in  $(0, 1)$  of  $1 - y = \exp\{-(1+\epsilon)y\}$ , and **whp** the largest component in supercritical  $G(n, p)$  is of order  $(1 + o(1))yn$  (see, for example, Remark 1.2 of [28]).

A fairly well-studied family of graphs are (percolation on) random  $d$ -regular graphs. For fixed  $d \geq 3$ , Goerdts [22] (see also the results of Alon, Benjamini, and Stacey [3]) showed that for a typical random  $d$ -regular graph  $G$  on  $n$  vertices,  $G_p$  exhibits a phase transition with respect to the size of its largest component around  $p = \frac{1}{d-1}$  (see also [26] for behaviour at criticality). In a subsequent work, Pittel [32] further refined this result: when  $p = \frac{1-\epsilon}{d-1}$ , for some small constant  $\epsilon > 0$ , **whp**<sup>2</sup> all components of  $G_p$  are of order at most  $O\left(\frac{\log n}{\epsilon^2}\right)$ , whereas when  $p = \frac{1+\epsilon}{d-1}$ , **whp** there is a unique giant component of order  $(1 + o_n(1))yn$  in  $G_p$  (where  $y = y(p)$  is defined according to (2)), and typically all other components of  $G_p$  are of order  $O(\log^{1+o(1)} n)$ . In the barely supercritical regime, that is when  $\epsilon = \epsilon(n) \rightarrow 0$  and  $\epsilon \gg n^{-1/3}$ , more precise results are known — see [31, 33]. There have also been several results considering percolation on graphs chosen uniformly at random from all graphs with a given degree sequence, see [9, 12, 18, 19, 24, 30, 33].

Another family of graphs which has been studied are high-girth expanders. For fixed  $d \geq 3$ , Alon, Benjamini, and Stacey [3] argued that when  $G$  is a  $d$ -regular high-girth expander on  $n$  vertices, then  $G_p$  exhibits a phase transition with respect to the size of its largest component around  $p = \frac{1}{d-1}$ : setting  $p = \frac{\lambda}{d-1}$ , when  $\lambda < 1$  **whp** all components of  $G_p$  are of order at most  $O(\log n)$ , whereas when  $\lambda > 1$  **whp** there is a *unique* giant component of order  $\Theta(n)$ . In a subsequent work, Krivelevich, Lubetzky, and Sudakov [28] showed that when  $\lambda > 1$ , for every  $\delta, b > 0$  there exists  $R$  such that if  $G$  is a  $d$ -regular graph with girth at least  $R$  and satisfies that for every  $U \subseteq V(G)$  with  $|U| \leq \frac{n}{2}$ ,  $e_G(U, U^c) \geq b|U|$ , then **whp**  $|1 - |L_1|/yn| \leq \delta$ , where  $y = y(p)$  is defined according to (2) (see also [2] for further generalisations and [34] for a related result). In terms of the typical sizes of the other components in the supercritical regime, Alon, Benjamini, and Stacey [3] showed that in the supercritical regime, the second largest component is of order  $O(n^a)$ , for some  $a = a(b) \in [0, 1)$ . Krivelevich, Lubetzky, and Sudakov [28] showed that this is in fact tight: for any  $a < 1$ , there are constant degree high-girth expanders  $G$ , where the second-largest component of supercritical  $G_p$  is in fact of order at least  $n^a$  for any  $a < 1$ .

Our main result gives sufficient conditions on a  $d$ -regular graph  $G$ , where  $d \geq 3$  is fixed, such that when  $p = \frac{\lambda}{d-1}$ , for some constant  $1 < \lambda < d-1$ , there typically emerges a unique giant component, and all other components are of order at most logarithmic in  $|V(G)|$ .

---

<sup>2</sup>With high probability, that is, with probability tending to one as  $n$  tends to infinity.

**Theorem 1.** Fix an integer  $d \geq 3$ . Let  $1 < \lambda < d - 1$  be a constant, and let  $p = \frac{\lambda}{d-1}$ . Let  $b, \alpha > 0$  be constants. Then, there exist constants  $c := c(\lambda, \alpha) > 0$  and  $C := C(d, \lambda, b) > 0$  such that the following holds. Let  $G$  be a  $d$ -regular graph on  $n$  vertices, satisfying:

- (P1) for every  $U \subseteq V(G)$  with  $|U| \leq n/2$ ,  $e_G(U, U^c) \geq b|U|$ ; and,
- (P2) for every  $U \subseteq V(G)$  with  $|U| \leq \log^C n$ ,  $e_G(U) \leq (1 + c)|U|$ .

Then, **whp**  $G_p$  contains a unique giant component,  $L_1$ , satisfying  $\left|1 - \frac{|L_1|}{yn}\right| \leq \alpha$ , where  $y$  is defined according to (2). Furthermore, **whp** every other component in  $G_p$  is of order at most  $5\lambda \log n / (\lambda - 1)^2$ .

Some comments are in place. We note that Assumption (P1) can be guaranteed by proving that the second largest eigenvalue of the adjacency matrix of  $G$  is bounded away from  $d$  (see [4]). Further, while we treat  $b$  as a constant, it follows readily from the proof that one can allow  $b$  to be  $1/\text{polylog}(n)$  by taking  $C$  to be a sufficiently large constant. The aforementioned construction of Krivelevich, Lubetzky, and Sudakov [28] demonstrates that *some* ‘local’ requirement is indeed necessary in order to guarantee that the second largest component is typically of order  $O(\log n)$ . In fact, it suffices to require Assumption (P2) only for connected sets  $U \subseteq V(G)$ . Observe that any graph  $G$  in which every two cycles of length at most  $C \log \log n$  are at distance at least  $C \log \log n$  satisfies Assumption (P2) as well. Since a random  $d$ -regular graph typically satisfies this (see, for example, [35]), and since a random  $d$ -regular graph **whp** satisfies Assumption (P1) (see [7]), our results apply to typical random  $d$ -regular graphs, and can be seen as a generalisation (and improvement) of the results of Pittel in the strictly supercritical regime [32]. It also follows that expanding  $d$ -regular  $n$ -vertex graphs  $G$  with girth  $\Omega(\log \log n)$  have the behaviour in the supercritical regime postulated by Theorem 1. Moreover, note that when  $\lambda = 1 + \epsilon$  for sufficiently small constant  $\epsilon > 0$ , we obtain that typically the second largest component is of order  $O(\log n / \epsilon^2)$ , similarly to the case of supercritical  $G(n, p)$ . Moreover, when  $d$  and  $\lambda$  are sufficiently large, we have that **whp** the second largest component is of order  $O(\log n / \lambda)$ , similarly to the ‘very’ supercritical regime of  $G(n, p)$  with  $p = \frac{\lambda}{n}$ , where the second largest component is typically of order  $O(\log n / (\lambda - 1 - \log \lambda))$ . While we treat  $\lambda > 1$  as a constant, with minor modifications our results also apply to parts of the barely supercritical regime, where  $\lambda = 1 + \epsilon$  and  $\epsilon = \epsilon(n) \rightarrow 0$ .

The paper is structured as follows. In Section 2 we set out some notation, a modification of the Breadth First Search (BFS) algorithm, and a couple of lemmas which we will use throughout the paper. Section 3 is devoted to the proof of our main result. Finally, in Section 4 we discuss our results, possible generalisations, and avenues for future research.

## 2 Preliminaries

Consider a graph  $H$  and a vertex  $v \in V(H)$ . Let  $C_H(v)$  be the connected component in  $H$  containing  $v$ . For an integer  $r$ , let  $B_H(v, r)$  be the ball of radius  $r$  in  $H$  centred at  $v$ , that is, the set of all vertices of  $H$  at a distance of at most  $r$  from  $v$ . For  $u, v \in V(H)$ , denote the distance in  $H$  between  $u$  and  $v$  by  $\text{dist}_H(u, v)$ . For  $S \subseteq V(H)$ , set  $\text{dist}_H(u, S) := \min_{v \in S} \text{dist}_H(u, v)$ . As is fairly standard,  $E_H(S, S^c)$  represents the set of edges in  $H$  with one endpoint in  $S$  and the other in  $V(H) \setminus S$ , and  $E_H(S)$  denotes the set of edges in the induced subgraph  $H[S]$ . We use  $e_H(S, S^c) := |E_H(S, S^c)|$  and  $e_H(S) := |E_H(S)|$ . We denote by  $N_H(S)$  the external neighbourhood of  $S$  in  $H$ . When the graph is clear from the context, the subscript may be omitted. All the logarithms are with the natural base. Throughout the paper, we systematically ignore rounding signs for the sake of clarity of presentation.

### 2.1 A modified Breadth First Search process

We will utilise the following modification of the classical Breadth First Search (BFS) exploration algorithm. The algorithm receives as input a graph  $H = (V, E)$  with an ordering  $\sigma$  on its vertices, a subset  $U \subseteq V$ , and a sequence  $(X_i)_{i=1}^{|E|}$  of i.i.d Bernoulli( $p$ ) random variables.

We maintain three sets throughout the process:  $S$ , the set of vertices whose exploration has been completed;  $Q$ , the set of vertices currently being explored, kept in a queue (first-in-first-out discipline); and  $T$ , the set of vertices which have yet been processed. We initialise  $S = \emptyset$ ,  $Q = U$ , and  $T = V \setminus U$ . The process stops once  $Q$  is empty.

At round  $t$  of the algorithm execution, we consider the first vertex  $v$  in  $Q$ , and query the first edge between  $v$  and  $T$  according to the order  $\sigma$  (that is, the edge connecting  $v$  to the first – according to  $\sigma$  – vertex in  $T$ ). If  $X_t = 1$ , we retain this edge and move its endpoint vertex in  $T$  to  $Q$ . If  $X_t = 0$ , we discard the edge and continue. If there are no neighbours of  $v$  in  $T$  left to query, we move  $v$  from  $Q$  to  $S$  and continue.

Note that once  $Q$  is empty,  $H_p[S]$  has the same distribution as  $\bigcup_{u \in U} C_{H_p}(u)$ . Moreover, at every moment we have queried (and answered in the negative) all edges between current  $S$  and  $T$ .

## 2.2 Auxiliary Lemmas

We will make use of the following two fairly standard probability bounds. The first one is a Chernoff-type tail bound for the binomial distribution (see, for example, Appendix A in [5]).

**Lemma 2.1.** *Let  $n \in \mathbb{N}$ , let  $p \in [0, 1]$ , and let  $X \sim \text{Bin}(n, p)$ . Then for any  $0 < t \leq \frac{np}{2}$ ,*

$$\mathbb{P}[|X - np| \geq t] \leq 2 \exp \left\{ -\frac{t^2}{3np} \right\}.$$

We will often use Lemma 2.1 when estimating a one-sided tail bound of  $X \sim \text{Bin}(n, p)$ , noting that

$$\mathbb{P}[X \leq a] \leq \mathbb{P}[|X - np| \geq a - np], \quad \text{and} \quad \mathbb{P}[X \geq a] \leq \mathbb{P}[|X - np| \geq a - np].$$

The second one is a variant of the well-known Azuma-Hoeffding inequality (see, for example, Chapter 7 in [5]).

**Lemma 2.2.** *Let  $m \in \mathbb{N}$  and let  $p \in [0, 1]$ . Let  $X = (X_1, X_2, \dots, X_m)$  be a random vector with range  $\Lambda = \{0, 1\}^m$  with each  $X_\ell$  distributed according to independent  $\text{Bernoulli}(p)$ . Let  $f : \Lambda \rightarrow \mathbb{R}$  be such that there exists  $C \in \mathbb{R}$  such that for every  $x, x' \in \Lambda$  which differ only in one coordinate,*

$$|f(x) - f(x')| \leq C.$$

*Then, for every  $t \geq 0$ ,*

$$\mathbb{P}[|f(X) - \mathbb{E}[f(X)]| \geq t] \leq 2 \exp \left\{ -\frac{t^2}{2C^2 m} \right\}.$$

We will also use the following lemma, relating our local density assumption to local vertex expansion.

**Lemma 2.3.** *Let  $d \geq 3$  and let  $G$  be a  $d$ -regular graph on  $n$  vertices satisfying Assumption (P2) with  $C > 1$  and  $0 < c < \frac{1}{4}$ . Then, for every  $v \in V(G)$ , we have  $|B(v, 2 \log \log n)| \geq \frac{10\lambda \log n}{(\lambda-1)^2}$ .*

*Proof.* Fix  $r \in \mathbb{N}$  and let  $U := B(v, r)$ . Note that  $U \cup N(U) = B(v, r+1)$ . If  $|U \cup N(U)| \leq \log^C n$ , then by (P2),  $e(U \cup N(U)) \leq (1+c)|U \cup N(U)|$ . On the other hand, again by (P2),  $e(U \cup N(U)) \geq d|U| - e(U) \geq (d-1-c)|U|$ . Hence,  $|U \cup N(U)| \geq \frac{|U|(d-1-c)}{1+c}$ . Thus, for every  $r \in \mathbb{N}$ ,

$$|B(v, r+1)| \geq \min \left\{ \log^C n, \left( 1 + \frac{d-1-c}{1+c} \right) |B(v, r)| \right\}.$$

Therefore, taking  $r_0 = \log_{1+\frac{d-1-c}{1+c}} \left( \frac{10\lambda \log n}{(\lambda-1)^2} \right)$ , we have that  $|B(v, r_0)| \geq \frac{10\lambda \log n}{(\lambda-1)^2}$ . All that is left is to note that  $r_0 \leq 2 \log \log n$ .  $\square$

Finally, we will utilise the following lemma, showing that if  $G$  is a  $d$ -regular graph satisfying Assumption (P2), then there are ‘many’ vertices in  $G$  which are far from any cycle. Formally,

**Lemma 2.4.** *Let  $d \geq 3$  and let  $G$  be a  $d$ -regular graph on  $n$  vertices satisfying Assumption (P2) with a sufficiently small constant  $c > 0$ . Then, there is a set  $X \subseteq V(G)$  with  $|X| \geq \left(1 - \frac{1}{(d-1)^{\frac{1}{16c}}}\right)n$  such that for every  $v \in X$ , there are no cycles in  $B\left(v, \frac{1}{16c}\right)$ .*

*Proof.* We first claim that in  $G$ , every two cycles of length at most  $\frac{1}{4c}$  are at distance at least  $\frac{1}{4c}$ . Otherwise, let  $U$  be the vertices of these two cycles, together with the vertices of a path of length at most  $\frac{1}{4c}$  connecting them. Then,  $|U| \leq 2 \cdot \frac{1}{4c} + \frac{1}{4c} = \frac{3}{4c}$  and  $e(U) \geq |U| + 1$  (since  $U$  is connected and contains at least two cycles). On the other hand, by (P1),

$$e(U) \leq (1+c)|U| \leq |U| + c \cdot \frac{3}{4c} < |U| + 1.$$

Let  $m$  be the number of cycles of length at most  $\frac{1}{4c}$  in  $G$ , denote them by  $C_1, \dots, C_m$ . For every  $i \in [m]$ , let  $N_i$  be the set of vertices in  $V(G)$  at distance at most  $\frac{1}{8c}$  from  $C_i$ . We then have that for every  $i \neq j \in [m]$ ,  $N_i \cap N_j = \emptyset$  (as otherwise we would have two cycles of length at most  $\frac{1}{4c}$  at distance at most  $2 \cdot \frac{1}{8c} = \frac{1}{4c}$ ). By the same reasoning, for every  $i \in [m]$ ,  $G[N_i]$  has only one cycle. We thus obtain that  $m \cdot (d-1)^{\frac{1}{8c}} \leq n$ , that is,  $m \leq \frac{n}{(d-1)^{1/(8c)}}$ .

Since  $m(d-1)^{1/(16c)} \leq \frac{n}{(d-1)^{1/(16c)}}$ , we conclude that there are at most  $\frac{n}{(d-1)^{1/(16c)}}$  vertices in  $G$  which are at distance at most  $\frac{1}{16c}$  from a cycle of length at most  $\frac{1}{4c}$ . Hence there is a set  $X \subseteq V(G)$ ,  $|X| \geq \left(1 - \frac{1}{(d-1)^{1/(16c)}}\right)n$ , such that every  $v \in X$  is at distance at least  $\frac{1}{16c}$  from any cycle of length at most  $\frac{1}{4c}$ . Thus, every  $v \in X$  satisfies that  $B\left(v, \frac{1}{16c}\right)$  contains no cycles.  $\square$

### 3 Proof of Theorem 1

Throughout the section, we assume  $C > 0$  is a large enough constant with respect to  $d, \lambda - 1$  and  $b$ , and that  $c > 0$  is a small enough constant with respect to  $\lambda - 1$  and  $\alpha$ .

We utilise a double-exposure/sprinkling argument á la Ajtai-Komlós-Szemerédi [1]. Let  $\delta := \delta(\lambda, d) > 0$  be a sufficiently small constant, satisfying that  $\lambda - \delta > 1$ . Let  $p_2 = \frac{\delta}{d-1}$  and let  $p_1$  be such that  $(1-p_1)(1-p_2) = 1-p$ , noting that  $p_1 \geq \frac{\lambda-\delta}{d-1}$  and that  $G_p$  has the same distribution as  $G_{p_1} \cup G_{p_2}$ . We abbreviate  $G_1 := G_{p_1}$  and  $G_2 := G_{p_1} \cup G_{p_2}$ . The overall strategy will be similar to that of the growing degree case [15], however the analysis itself (and where the difficulties lie) is quite different. We begin by considering ‘large’ components in  $G_1$ . We show that typically there are no components in  $G_1$  (nor in  $G_2$ ) whose size is between  $5\lambda \log n / (\lambda - 1)^2$  and  $\log^C n$  (Lemma 3.1). Here, unlike the growing degree case, this does not follow from a first moment argument, and we will utilise the BFS algorithm in order to estimate the probability a fixed vertex lies in a component of a given size. Utilising our modified BFS algorithm, we further show that **whp** in  $G_1$ , every vertex is within distance  $O(\log \log n)$  from a ‘large’ component (Lemma 3.2). Then, in a manner similar to that of [1], using these two properties, we show that typically after sprinkling with probability  $p_2$ , all ‘large’ components in  $G_1$  merge into a unique component  $L_1$  in  $G_2$  (Lemma 3.3). Finally, we show that when sprinkling with probability  $p_2$ , we typically do not ‘accidentally’ merge small components (of order  $O(\log n)$ ) in  $G_1$  into a large component (of order  $\Omega(\log^C n)$ ) in  $G_2$  (Lemma 3.4). All that is then left is to argue about the size of  $L_1$ , which we do by utilising Lemma 2.4, in a fashion somewhat similar to [28].

More formally, let  $W$  be the set of vertices in ‘large’ components in  $G_1$ , that is,

$$W := \left\{ v \in V(G) : |C_{G_1}(v)| \geq \frac{5\lambda \log n}{(\lambda - 1)^2} \right\}.$$

We begin by showing that **whp** there are no components in  $G_p$  whose order is between  $\frac{5\lambda \log n}{(\lambda-1)^2}$  and  $\log^C n$ .

**Lemma 3.1.** *Fix  $v \in V(G)$ , and let  $k \in [\frac{1}{16c}, \log^C n]$ . The probability that  $|C_{G_p}(v)| = k$  is at most  $\exp\left\{-\frac{(\lambda-1)^2 k}{4\lambda}\right\}$ . In particular, there are no components in  $G_p$  whose order is between  $\frac{5\lambda \log n}{(\lambda-1)^2}$  and  $\log^C n$ .*

*Proof.* We will utilise the BFS algorithm described in Section 2.1, with  $U = \{v\}$  (that is, we initialise  $Q = \{v\}$ ), with  $H = G$  and with probability  $p$ .

Fix  $k \in [1/(16c), \log^C n]$ . If  $|C_{G_p}(v)| = k$ , then there is some moment  $t$  where  $Q$  is empty and  $|S| = k$ . Since  $k \leq \log^C n$ , by Assumption (P2),  $e_G(S) \leq (1+c)k$ , and since  $G$  is  $d$ -regular and  $Q = \emptyset$ ,  $e_G(S, T) = e_G(S, S^c) \geq (d-2-2c)k$ . Hence, we have had at least  $(d-2-2c)k$  queries corresponding to  $e_G(S, T)$ , and at least additional  $k-1$  queries corresponding to the internal edges of  $C_{G_p}(v)$  which have been explored. Further, only  $k-1$  queries were answered in the positive. Therefore, by Lemma 2.1,

$$\begin{aligned} \mathbb{P}[|C_{G_p}(v)|] &\leq \mathbb{P}\left[\text{Bin}\left(k(d-1-2c)-1, \frac{\lambda}{d-1}\right) \leq k-1\right] \\ &\leq \exp\left\{-\frac{(\lambda-1)^2 k^2}{4\lambda k}\right\} = \exp\left\{-\frac{(\lambda-1)^2 k}{4\lambda}\right\}, \end{aligned}$$

where we assumed that  $c$  is sufficiently small with respect to  $\lambda-1$ .

As for the second part of the lemma's statement, for  $k \in [5\lambda \log n/(\lambda-1)^2, \log^C n]$  the above implies the probability that a fixed  $v$  is in a component of order  $k$  is at most  $\exp\left\{-\frac{(\lambda-1)^2 k}{4\lambda}\right\} = o\left(\frac{1}{n \log^C n}\right)$ . Therefore, the union bound over the  $n$  possible choices of  $v$  and  $\log^C n$  choices of  $k$  completes the proof.  $\square$

Note that by our assumptions on  $p_1$ , the above holds (with the same proof) in  $G_1$ .

We now turn to show that typically every vertex is within distance  $O(\log \log n)$  from a vertex in  $W$ .

**Lemma 3.2.** *Whp, for every  $v \in V(G)$  we have that  $\text{dist}_G(u, W) \leq 2 \log \log n$ .*

*Proof.* Fix  $v \in V(G)$ . By Lemma 2.3, we have that  $|B(v, 2 \log \log n)| \geq \frac{10\lambda \log n}{(\lambda-1)^2}$ . Let  $Y_v$  be an arbitrary set of  $\frac{10\lambda \log n}{(\lambda-1)^2}$  vertices in  $B(v, 2 \log \log n)$ . We now run the BFS algorithm described in Section 2.1 with  $U = Y_v$  (that is, we initialise  $Q = Y_v$ ),  $H = G$ , and probability  $p_1$ .

Suppose towards contradiction that at the moment  $t$  when  $Q$  emptied,  $|S| := s \leq \log^C n$ . Then, all the edges between  $S$  and  $T = V(G) \setminus S$  have been queried and answered in the negative. By Assumption (P2), we have that  $e_G(S, T) = e_G(S, S^c) \geq (d-2-2c)s$ . Hence,  $t \geq (d-2-2c)s + s - \frac{10\lambda \log n}{(\lambda-1)^2}$ , and we have received  $s - \frac{10\lambda \log n}{(\lambda-1)^2}$  positive answers. By Lemma 2.1, the probability of this event is at most

$$\mathbb{P}\left[\text{Bin}\left((d-1-2c)s - \frac{10\lambda \log n}{(\lambda-1)^2}, \frac{\lambda-\delta}{d-1}\right) \leq s - \frac{10\lambda \log n}{(\lambda-1)^2}\right] \leq \exp\left\{-\frac{((\lambda-1)s/2)^2}{4\lambda s}\right\} = o(1/n),$$

where we used our assumption that  $c, \delta$  are sufficiently small with respect to  $\lambda-1$ , and the fact that  $s \geq \frac{10\lambda \log n}{(\lambda-1)^2}$ .

Thus, by the union bound over the  $n$  possible choices of  $v$ , we have that **whp** for every  $v \in V(G)$ , there is a set  $Y_v \subseteq B(v, 2 \log \log n)$  of order  $\frac{10\lambda \log n}{(\lambda-1)^2}$ , such that

$$\left|\bigcup_{u \in Y_v} C_{G_1}(u)\right| \geq \log^C n.$$

Assuming that  $C > 3$ , we conclude that **whp** there is a component of order  $\Omega(\log^2 n) \geq \frac{5\lambda \log n}{(\lambda-1)^2}$  in  $G_1$  at distance at most  $2 \log \log n$  from every vertex  $v \in V(G)$ .  $\square$

We are now ready to argue that after sprinkling with probability  $p_2$ , **whp** all components in  $W$  merge.

**Lemma 3.3.** *Whp there is a component  $K$  in  $G_2$ , such that  $W \subseteq V(K)$ .*

*Proof.* By Lemma 3.1, **whp** there are no components in  $G_1$  whose size is between  $\frac{5\lambda \log n}{(\lambda-1)^2}$  and  $\log^C n$ . Thus, **whp** every component in  $G_1[W]$  is of order at least  $\log^C n$ . Further, by Lemma 3.2, **whp** every  $v \in V(G)$  is at distance at most  $2 \log \log n$  from some  $w \in W$ . We continue assuming these properties hold deterministically.

It suffices to show that **whp** for every partition of  $W$  into two  $G_1$ -component-respecting parts  $A$  and  $B$ , with  $a = |A| \leq |B|$ , there exists a path in  $G_{p_2}$  between  $A$  and  $B$ . To that end, let  $A'$  be  $A$  together with all the vertices in  $V(G) \setminus B$  which are at distance at most  $2 \log \log n$  from some vertex in  $A$ . Similarly, let  $B'$  be  $B$  together with all the vertices in  $V(G) \setminus A'$  which are at distance at most  $2 \log \log n$  from some vertex in  $B$ . Note that  $V(G) = A' \sqcup B'$ , and thus by (P1),  $e_G(A', B') = e_G(A', V(G) \setminus A') \geq b|A|$ . We can very crudely extend these edges into at least  $\frac{b \cdot a}{d^5 \log \log n}$  edge-disjoint paths (in  $G$ ) of length at most  $4 \log \log n$  between  $A$  and  $B$  (indeed, every edge belongs to at most  $d^{4 \log \log n} \cdot 4 \log \log n < d^{5 \log \log n}$  paths of length at most  $4 \log \log n$ ). Since every component of  $W$  is of size at least  $\log^C n$  and since  $W = A \sqcup B$  is a component-respecting partition, given that  $|A| = a$  there are at most  $\sum_{i=1}^{a/\log^C n} \binom{n/\log^C n}{a/\log^C n} \leq n^{a/\log^C n}$  ways to choose  $A$  (and hence the partition). Thus, by the union bound, the probability there exists such a partition without a path in  $G_{p_2}$  between  $A$  and  $B$  is at most

$$\begin{aligned} \sum_{a=\log^C n}^{n/2} n^{a/\log^C n} (1 - p_2^{4 \log \log n})^{\frac{b \cdot a}{d^5 \log \log n}} &\leq \sum_{a=\log^C n}^{n/2} n^{a/\log^C n} \exp \left\{ -ab (p_2/d)^{5 \log \log n} \right\} \\ &\leq \sum_{a=\log^C n}^{n/2} \exp \left\{ a \left( \frac{1}{\log^{C-1} n} - \frac{b}{\log^{10 \log d + 5 \log(1/\delta)} n} \right) \right\} \\ &\leq n \cdot \exp \left\{ -\log^C n \frac{b}{2 \log^{10 \log d + 5 \log(1/\delta)} n} \right\} = o(1), \end{aligned}$$

where we assumed that  $C$  is sufficiently large with respect to  $b, d$ , and  $\lambda - 1$ .  $\square$

By Lemma 3.1, **whp** there are no components in  $G_1$ , nor in  $G_2$ , whose order is between  $5\lambda \log n / (\lambda - 1)^2$  and  $\log^C n$ . Further, by Lemma 3.3 **whp** all components in  $G_1$  whose order was at least  $5\lambda \log n / (\lambda - 1)^2$  merged into a unique component. Note, however, that we still need to rule out the existence of components outside of  $W$  whose order is at least  $\log^C n$  in  $G_2$ : any such component would be composed of components in  $G_1$  whose size is at most  $5\lambda \log n / (\lambda - 1)^2$ . We give two proofs for this, showcasing different approaches and arguments.

**Lemma 3.4.** *Whp, there is no component in  $G_2$  of order at least  $\log^C n$ , which does not intersect  $W$ .*

*Proof.* We begin by exposing  $G_1$ . Let us show that after sprinkling edges with probability  $p_2$ , there is no component in  $G_2$  of order at least  $\log^C n$  which does not intersect  $W$ .

Suppose first that  $G_1$  is such that one can create a connected vertex set  $M$  satisfying  $|M| \in [\log^2 n, 2 \log^2 n]$ ,  $M \cap W = \emptyset$  and  $E_{G_1}[M, V(G) \setminus M] = \emptyset$  by adding at most  $t = \frac{(\lambda-1)^2 \log^2 n}{5\lambda \log((d-1)/\delta)}$  edges to  $G_1$ . Then, the probability that  $G_2$  contains a connected component  $K$  whose size is in the interval  $[\log^2 n, 2 \log^2 n]$  is at least

$$\begin{aligned} p_2^t (1 - p_2)^{d \cdot 2 \log^2 n} &\geq \exp \left\{ -t \log \left( \frac{d-1}{\delta} \right) - \delta \cdot 4 \log^2 n \right\} \\ &\geq \exp \left\{ -\frac{((\lambda-1)^2 + 20\delta\lambda) \log^2 n}{5\lambda} \right\}. \end{aligned}$$

On the other hand, by Lemma 3.1, the probability there is a connected component in  $G_2$  whose order lies in the interval  $[\log^2 n, 2 \log^2 n]$  is at most

$$n \cdot 2 \log^2 n \cdot \exp \left\{ -\frac{(\lambda - 1)^2 \log^2 n}{4\lambda} \right\},$$

which is a contradiction since  $\frac{(\lambda-1)^2 \log^2 n}{4\lambda} - 2 > \frac{((\lambda-1)^2 + 20\delta\lambda) \log^2 n}{5\lambda}$ , where we assumed that  $\delta$  is sufficiently small with respect to  $\lambda - 1$ .

We can thus assume that in  $G_1$ , in order to create a connected set  $M$  satisfying  $|M| \in [\log^2 n, 2 \log^2 n]$ ,  $M \cap W = \emptyset$  and  $E_{G_1}[M, V(G) \setminus M] = \emptyset$ , we must add at least  $t$  edges. Utilising that, let us now show that **whp** when exploring the connected component in  $G_2$  of any component  $K$  of  $G_1[V(G) \setminus W]$ , we could not uncover a component of order at least  $\log^C n$  which does not intersect  $W$ .

To that end, consider the following variant of the BFS algorithm, which receives as input the graph  $G$  with an order  $\sigma$  on its vertices, the subgraph  $G_1 \subseteq G$ , a component  $K$  in  $G_1[V(G) \setminus W]$ , and a sequence  $(X_i)_{i=1}^{nd/2}$  of independent Bernoulli( $p_2$ ) random variables. As in Section 2.1, the algorithm maintains three sets:  $S$ , the set of vertices whose exploration has been completed;  $Q$ , the set of vertices currently being explored, kept in a queue (first-in-first-out discipline); and  $T$ , the set of vertices which have yet been processed. We initialise  $S = \emptyset$ ,  $Q = V(K)$ , and  $T = V \setminus (K \cup W)$ . The process stops once  $Q$  is empty.

At round  $\tau$ , we consider the first vertex  $v$  in  $Q$ , and query the first edge from  $v$  to  $T$  according to the order  $\sigma$ , denote its endpoint in  $T$  by  $u$ . If  $X_\tau = 1$ , we retain this edge and move the set  $C_{G_1}(u)$  from  $T$  to  $Q$ . If  $X_t = 0$ , we discard the edge and continue. If there are no neighbours of  $v$  in  $T$  left to query, we move  $v$  from  $Q$  to  $S$  and continue.

Note that at the end of this process, the component of  $G_2[V(G) \setminus W]$  intersecting with  $K$  has the same distribution as  $G_2[S]$ . If the component of  $G_2[V(G) \setminus W]$  intersecting with  $K$  is of order at least  $\log^C n$ , then must have been a moment  $\tau$  where  $|S \cup Q| \in [\log^2 n, 2 \log^2 n]$ . Indeed, every component of  $G_1[V(G) \setminus W]$  is of size at most  $5\lambda \log n / (\lambda - 1)^2$ , and at the first time when  $|S \cup Q| \geq \log^2 n$ , we have that  $|S \cup Q| \leq \log^2 n + 5\lambda \log n / (\lambda - 1)^2 < 2 \log^2 n$ . Thus,  $\tau \leq 2d \log^2 n$ . On the other hand, by our assumption, by moment  $\tau$  we received at least  $t$  positive answers. By the union bound over the at most  $n$  possible choices of  $K$ , and by Lemma 2.1, the probability of this event is at most

$$n \cdot \mathbb{P} \left[ \text{Bin} \left( 2d \log^2 n, \frac{\delta}{d-1} \right) \geq \frac{(\lambda - 1)^2 \log^2 n}{5\lambda \log((d-1)/\delta)} \right] \leq n \cdot 2 \exp \{ -\Omega(\log^2 n) \} = o(1),$$

where we assumed that  $\delta$  is small enough with respect to  $\lambda - 1$  and  $d$ .  $\square$

Let us also give a second proof, which is shorter, and somewhat similar to the BFS-type arguments given previously in this paper.

*Alternative proof of Lemma 3.4.* For  $v \in V(G)$ , let  $\mathcal{A}_v$  be the event that  $C_{G_2}(v) \cap W = \emptyset$  and  $|C_{G_2}(v)| \geq \log^2 n$ . By Lemma 3.1, the probability that there is a component whose size is at least  $\log^2 n$  but less than  $\log^C n$  is  $o(1)$ . Thus, we have that the probability of an event violating the statement of the lemma is at most  $\mathbb{P} \left[ \bigcup_{v \in V(G)} \mathcal{A}_v \right] + o(1)$ .

Let  $(X_i)_{i=1}^{|E|}$  be a sequence of i.i.d Bernoulli( $p_1$ ) random variables, and let  $(Y_i)_{i=1}^{|E|}$  be a sequence of i.i.d. Bernoulli ( $p_2$ ) random variables. We will run a BFS-type algorithm, very similar to the one in Section 2.1, however, we will utilise both sequences of random variables.

Fix  $v \in V(G)$ . We initialise  $Q = \{v\}$ . Now, as long as  $|S \cup Q| \leq \log^2 n$ , when we query the  $t$ -th edge, we retain it if  $X_t = 1$  or  $Y_t = 1$ , and discard it only if  $X_t = Y_t = 0$ . Once  $|S \cup Q| = \log^2 n$  (and at any subsequent moment), when we query the  $t$ -th edge, we retain it if  $X_t = 1$ , and discard it if  $X_t = 0$ . Note that in this manner, up until  $|S \cup Q| = \log^2 n$ , the exploration of the component of  $v$  is in  $G_2$ , and afterwards we continue the exploration in  $G_1$ . Also, each query obtains a positive answer independently and with probability at least  $p_1$ .

Note that if  $\mathcal{A}_v$  occurs, then we must have reached a moment where  $|S \cup Q| = \log^2 n$ . Suppose that  $C_{G_2}(v) \cap W = \emptyset$ . Then, the above process must have ended before  $|S \cup Q| = \log^4 n$ . Indeed, otherwise, there are vertices in  $C_{G_2}(v)$  belonging to components of size at least  $\log^4 n / \log^2 n = \log^2 n$  in  $G_1$ , and thus (by definition) intersect with  $W$ . Hence, the probability that  $\mathcal{A}_v$  occurs is at most the probability that the above process stopped at some moment  $t$  where  $\log^2 n \leq |S| \leq \log^4 n$ . Let  $k := |S|$ . At that moment, we had exactly  $k-1$  positive answers in the algorithm's run. Moreover, assuming that  $C \geq 4$ , by Property (P2) we had additional  $e(S, T) = e(S, S^c) \geq (d-2-2c)k$  queries, all answered in the negative. Hence, we had at least  $(d-1-2c)k-1$  queries and received only  $k-1$  positive answers. Noting that  $((d-1-2c)k-1) \cdot p_1 > k-1$ , we have by Lemma 2.1 that

$$\begin{aligned} \mathbb{P}[\mathcal{A}_v] &\leq \mathbb{P}[\text{Bin}((d-1-2c)k-1, p_1) \leq k-1] \\ &\leq \exp\{-\Theta(k)\} = o(1/n), \end{aligned}$$

where we used that  $k \geq \log^2 n$ . Thus, by the union bound,  $\mathbb{P}\left[\bigcup_{v \in V(G)} \mathcal{A}_v\right] = o(1)$ , completing the proof.  $\square$

We are now ready to prove Theorem 1.

*Proof of Theorem 1.* By Lemma 3.3, **whp** there is a unique component  $L_1$  in  $G_2$ , such that  $W \subseteq V(L_1)$ . By Lemma 3.1, **whp** any component in  $G_2$  besides  $L_1$  is either of size at most  $5\lambda \log n / (\lambda-1)^2$ , or of size at least  $\log^C n$ . By Lemma 3.4, **whp** any component in  $G_2$  whose size is at least  $\log^C n$  intersects with  $W$ , and is thus part of  $L_1$ . Thus, **whp** all components of  $G_2$  besides  $L_1$  are of order at most  $5\lambda \log n / (\lambda-1)^2$ .

Let us now show that **whp**  $\left|1 - \frac{|L_1|}{yn}\right| \leq \alpha$ , where  $y = y(p)$  is defined according to (1).

The probability a vertex belongs to a component of order at least  $\log^C n$  in  $G_p$  is stochastically dominated by the probability that the root of an infinite  $d$ -regular tree belongs to an infinite cluster after  $p$ -bond-percolation. Thus, by standard results on Galton-Watson processes (see, for example, [16, Chapter 4.3])  $|L_1| \leq (1 + o(1))yn$ .

Let  $Z_1$  be the random variable counting the number of vertices in components of order at least  $\frac{1}{16c}$  in  $G_2$ . By Lemma 2.4, there exists a set  $X \subseteq V(G)$ ,  $|X| \geq \left(1 - \frac{1}{(d-1)\frac{1}{16c}}\right)n$  such that for every  $v \in X$ , there are no cycles in  $B(v, \frac{1}{16c})$ . Hence, by standard results [16, Chapter 4.3], we have that for every  $v \in X$ ,  $\mathbb{P}[|C_{G_2}(v)| \geq \frac{1}{16c}] \geq (1 - o_c(1))y$ , where  $o_c(1)$  tends to zero as  $c$  tends to zero. Thus,  $\mathbb{E}[|Z_1|] \geq (1 + o_c(1))y \cdot \left(1 - \frac{1}{(d-1)\frac{1}{16c}}\right)n = (1 - o_c(1))yn$ . To show that  $|Z_1|$  is well concentrated around its mean, consider the standard edge-exposure martingale. Every edge can change the value of  $|Z_1|$  by at most  $\frac{1}{8c}$ . Hence, by Lemma 2.2,

$$\mathbb{P}\left[||Z_1| - \mathbb{E}[|Z_1|]| \geq n^{2/3}\right] \leq 2 \exp\left\{-\frac{n^{4/3}}{2 \cdot \frac{nd}{2}} \cdot \frac{1}{64c^2}\right\} = o(1).$$

Therefore, **whp**,  $|Z_1| \geq (1 - o_c(1))yn$ .

Let  $Z_2$  be the random variable counting the number of vertices in components of  $G_2$  whose order lies in the interval  $[\frac{1}{16c}, \log^C n]$ . By Lemma 3.1, the probability  $v \in V(G)$  belongs to such a component is at most  $\sum_{k=1}^{\log^C n} \exp\left\{-\frac{(\lambda-1)^2 k}{4}\right\} \leq \exp\left\{-\frac{(\lambda-1)^2}{70\lambda \cdot c}\right\}$ . Thus  $\mathbb{E}[|Z_2|] \leq o_c(1)n$ , where we assumed that  $c$  is sufficiently small with respect to  $\lambda-1$ . Once again, let us consider the standard edge-exposure martingale. Every edge can change the value of  $|Z_2|$  by at most  $2\log^C n$ . Hence, by Lemma 2.2,

$$\mathbb{P}\left[||Z_2| - \mathbb{E}[|Z_2|]| \geq n^{2/3}\right] \leq 2 \exp\left\{-\frac{n^{4/3}}{2 \cdot \frac{nd}{2} \cdot 4 \log^{2C} n}\right\} = o(1).$$

Therefore, we obtain that the number of vertices in components of order at least  $\log^C n$  in  $G_2$  is **whp**  $(1 - o_c(1))yn - (1 + o(1))o_c(1)n$ . Thus, given  $\alpha$  we can choose  $c$  small enough such that **whp** there are at least  $(y - \alpha)n$  vertices in components of order at least  $\log^C n$  in  $G_2$ . By Lemma 3.4 every component of order  $\log^C n$  in  $G_2$  intersects with  $W$ . Thus, the number of vertices in components that intersect with  $W$  is at least  $(y - \alpha)n$ . By Lemma 3.3, all the vertices in  $W$  merge into a unique component  $L_1$ , and hence **whp**  $|L_1| \geq (y - \alpha)n$ . Altogether, we have that  $\left|1 - \frac{|L_1|}{yn}\right| \leq \alpha$ .  $\square$

## 4 Discussion

We showed that for a fixed  $d \geq 3$ , any  $d$ -regular  $n$ -vertex graph  $G$  which satisfies a fairly mild global expansion assumption (P1), and does not have dense sets of polylogarithmic order (Property (P2)), exhibits a phase transition around  $p = \frac{1}{d-1}$  similar to that of the binomial random graph  $G(n, p)$  around  $p = \frac{1}{n}$  (and alike to that of percolation on a random  $d$ -regular graph) — that is, the typical emergence of a unique giant component  $L_1$  of order linear in  $n$ , while **whp** all other components are of order at most logarithmic in  $n$ .

With slight adaptation, it follows from the above proof that for a fixed  $d \geq 3$ , given an  $n$ -vertex graph  $G$  with *minimum degree*  $d$ , when  $p = \frac{\lambda}{d-1}$  with  $\lambda > 1$ , **whp**  $G_p$  contains a giant component of order linear in  $n$ , and all the other components are of order logarithmic in  $n$ . Indeed, if we rely on the alternative proof given for Lemma 3.4, the only place where we use an upper bound on the maximum degree of  $G$  is in Lemma 3.3 (when estimating the number of edge-disjoint paths between  $A$  and  $B$ ). To overcome that, one can take the more delicate approach as given in [15]. First, we show that **whp** for every  $v \in V(G)$ , a constant fraction of the vertices in  $B(v, 2\log \log n)$  are in  $W$ . Using that, we can show that given a  $G_1$ -component-respecting partition  $A \sqcup B$  of  $W$ , **whp** we can construct an  $\Omega(d)$ -regular tree of depth at most  $2\log \log n$  from every  $v \notin W$ , whose leaves are either in  $A$  or in  $B$ . Then, when sprinkling with probability  $p_2$  on these trees, we can show that **whp** for any such partition, there is a path between  $A$  and  $B$ .

We note that while one might anticipate the same could hold if we consider  $p \geq (1 + \epsilon)p_c(G)$ , where  $p_c(G)$  is the critical probability of the graph  $G$  and  $\epsilon > 0$  is a small constant, this is quite false. Indeed, consider the graph  $G$  constructed as follows (where we note that the construction is similar to the construction given in [28]). Let  $0 < a < 1$  be a constant, and let  $H$  be a  $100d$ -regular graph on  $n - \log n \cdot n^a$  vertices, which is a fairly good expander and is locally sparse (for example, a random  $100d$ -regular graph on  $n - \log n \cdot n^a$  typically satisfies this). Let  $T$  be a  $C'd$ -regular tree (for some sufficiently large  $C' > 0$ ) of depth  $a \log_{C'd-1} n/2$ , and from each leaf of  $T$ , let us grow a  $d$ -regular tree on about  $n^{a/2}$  vertices. Denote by  $T'$  the obtained graph, and note that it has  $\approx n^{a/2} \cdot (C'd - 1)^{a \log_{C'd-1} n/2} = n^a$  vertices. We now take  $\log n$  copies of  $T'$ , and identify the leaves of each copy with distinct vertices in  $H$ . Let  $G$  be the obtained graph. Note that  $G$  is also a good expander and locally sparse. We then have that  $p_c(G) = \frac{1+o(1)}{100(d-1)}$ , and when  $p = (1 + \epsilon)p_c(G)$ , for large enough  $C'$  the second-largest component of  $G_p$  would be of order  $\Omega(n^{a/2})$ .

As mentioned in the introduction, Krivelevich, Lubetzky, and Sudakov [28] showed the existence of  $d$ -regular graphs on  $n$  vertices, satisfying Assumption (P1) and with large (yet constant) girth, such that the second largest component in the supercritical regime is typically of order polynomial in  $n$ . Thus, while it is clear that some additional requirement (to (P1)) is necessary in order to ensure that  $G$  exhibits a phase transition similar to that of  $G(n, p)$ , it remains open whether Assumption (P2) is (essentially) tight, or could it be relaxed — can we ensure that  $G$  exhibits a phase transition similar to that of  $G(n, p)$ , replacing our local density assumption with one that holds for sets up to size, say,  $O(\log n)$ ? Furthermore, as mentioned in the discussion after Theorem 1, it suffices for  $G$  to satisfy (P1) and have girth of order  $\Omega(\log \log n)$  in order to ensure such a phase transition. Can this assumption on the girth of  $G$  be further relaxed? What assumption on the girth of  $G$  would suffice?

## References

- [1] M. Ajtai, J. Komlós, and E. Szemerédi. Largest random component of a  $k$ -cube. *Combinatorica*, 2(1):1–7, 1982.
- [2] Y. Alimohammadi, C. Borgs, and A. Saberi. Locality of random digraphs on expanders. *Ann. Probab.*, 51(4):1249–1297, 2023.
- [3] N. Alon, I. Benjamini, and A. Stacey. Percolation on finite graphs and isoperimetric inequalities. *Ann. Probab.*, 32(3):1727–1745, 2004.
- [4] N. Alon and V. D. Milman.  $\lambda_1$ , isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B*, 38(1):73–88, 1985.
- [5] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.
- [6] K. B. Athreya and P. E. Ney. *Branching processes*, volume 196 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York-Heidelberg, 1972.
- [7] B. Bollobás. The isoperimetric number of random regular graphs. *European J. Combin.*, 9(3):241–244, 1988.
- [8] B. Bollobás. *Random graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition, 2001.
- [9] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Percolation on dense graph sequences. *Ann. Probab.*, 38(1):150–183, 2010.
- [10] B. Bollobás, Y. Kohayakawa, and T. Łuczak. The evolution of random subgraphs of the cube. *Random Structures Algorithms*, 3(1):55–90, 1992.
- [11] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, Cambridge, 2006.
- [12] B. Bollobás and O. Riordan. An old approach to the giant component problem. *J. Comb. Theory, Ser. B*, 113:236–260, 2015.
- [13] M. Collares, J. Doolittle, and J. Erde. The evolution of the permutahedron. *arXiv preprint arXiv:2404.17260*, 2024.
- [14] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Percolation on high-dimensional product graphs. *Random Structures Algorithms*, 66(1):Paper No. e21268, 10, 2025.
- [15] S. Diskin and M. Krivelevich. Components, large and small, as they should be I: supercritical percolation in regular graphs of growing degree. *arXiv:2408.04597*, 2024.
- [16] R. Durrett. *Probability: theory and examples*. Cambridge University Press, Cambridge, 2019.
- [17] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.
- [18] N. Fountoulakis. Percolation on sparse random graphs with given degree sequence. *Internet Math.*, 4(4):329–356, 2007.
- [19] N. Fountoulakis, F. Joos, and G. Perarnau. Percolation on random graphs with a fixed degree sequence. *SIAM J. Discrete Math.*, 36(1):1–46, 2022.

- [20] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.
- [21] A. Frieze, M. Krivelevich, and R. Martin. The emergence of a giant component in random subgraphs of pseudo-random graphs. *Random Structures Algorithms*, 24(1):42–50, 2004.
- [22] A. Goerdts. The giant component threshold for random regular graphs with edge faults. *Theoret. Comput. Sci.*, 259(1-2):307–321, 2001.
- [23] G. Grimmett. *Percolation*. Springer, Berlin, 1999.
- [24] S. Janson. On percolation in random graphs with given vertex degrees. *Electron. J. Probab.*, 14:86–118, 2009.
- [25] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, 2000.
- [26] F. Joos and G. Perarnau. Critical percolation on random regular graphs. *Proc. Amer. Math. Soc.*, 146(8):3321–3332, 2018.
- [27] H. Kesten. *Percolation theory for mathematicians*. Birkhäuser, Boston, MA, 1982.
- [28] M. Krivelevich, E. Lubetzky, and B. Sudakov. Asymptotics in percolation on high-girth expanders. *Random Structures Algorithms*, 56(4):927–947, 2020.
- [29] L. Lichev. The giant component after percolation of product graphs. *J. Graph Theory*, 99(4):651–670, 2022.
- [30] L. Lichev, D. Mitsche, and G. Perarnau. Percolation on dense random graphs with given degrees. *J. Comb. Theory, Ser. B*, 167:250–282, 2024.
- [31] A. Nachmias and Y. Peres. Critical percolation on random regular graphs. *Random Structures Algorithms*, 36(2):111–148, 2010.
- [32] B. Pittel. Edge percolation on a random regular graph of low degree. *Ann. Probab.*, 36(4):1359–1389, 2008.
- [33] O. Riordan. The phase transition in the configuration model. *Combin. Probab. Comput.*, 21(1-2):265–299, 2012.
- [34] R. van der Hofstad. The giant in random graphs is almost local. *arXiv:2103.11733*, 2023.
- [35] N. C. Wormald. Models of random regular graphs. In *Surveys in Combinatorics, 1999*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.