CRITICAL PERCOLATION ON RANDOM REGULAR GRAPHS

FELIX JOOS AND GUILLEM PERARNAU

Abstract. We show that for all $d \in \{3, \dots, n-1\}$ the size of the largest component of a random d-regular graph on n vertices around the percolation threshold $p = 1/(d-1)$ is $\Theta(n^{2/3})$, with high probability. This extends known results for fixed $d \geq 3$ and for $d = n-1$, confirming a prediction of Nachmias and Peres on a question of Benjamini. As a corollary, for the largest component of the percolated random d-regular graph, we also determine the diameter and the mixing time of the lazy random walk. In contrast to previous approaches, our proof is based on a simple application of the switching method.

1. Introduction

For every $d \in \{3, \dots, n-1\}$, let $\mathcal{G}_{n,d}$ be the set of all simple and vertex-labelled d-regular graphs on n vertices and let $G_{n,d}$ be a graph chosen uniformly at random from $\mathcal{G}_{n,d}$. For $p \in [0,1]$, let $G_{n,d,p}$ be a graph obtained from $G_{n,d}$ by retaining each edge independently with probability p. The goal of this paper is to study the order of the largest component of $G_{n,d,p}$, denoted by $L_1(G_{n,d,p})$, in terms of n, d and p.

Most of the literature in the area focuses either on fixed $d \geq 3$ or on $d = n-1$. Goerdt [8] showed the existence of a critical probability, $p_{crit} := 1/(d-1)$, such that for every fixed $d \geq 3$ and every $\epsilon > 0$ the following holds with probability $1 - o(1)$: if $p \leq (1-\epsilon)p_{crit}$, then $L_1(G_{n,d,p}) = O(\log n)$, while if $p \geq (1+\epsilon)p_{crit}$, then $L_1(G_{n,d,p}) = \Theta(n)$. Similar results were also obtained in a more general setting by Alon, Benjamini and Stacey [1]. For $d = n-1$, the random graph $G_{n,d,p}$ corresponds to the classic Erdős-Rényi random graph $G_{n,p}$. In their seminal paper [5], Erdős and Rényi proved that for every $\epsilon > 0$, the following holds with probability $1 - o(1)$: if $p \leq (1-\epsilon)/n$, then the largest component of $G_{n,p}$ has order $O(\log n)$, if $p = 1/n$ (critical probability), then it has order $\Theta(n^{2/3})$, while if $p \geq (1+\epsilon)/n$, then it has linear order.

Both for fixed $d \geq 3$ and for $d = n-1$, the behaviour around the critical probability has attracted a lot of interest. It is well established that the critical window in $G_{n,p}$ around $p = 1/n$ is of order $n^{-1/3}$ (see e.g. [21]). More precise estimates can be found in [14]. Benjamini posed the problem of determining the width of the critical window in $G_{n,d,p}$ around $p_{crit} = 1/(d-1)$ (see [20, 22]). Nachmias and Peres [20] and Pittel [22], independently showed that the critical window exhibits mean-field behaviour for fixed $d \geq 3$, namely, the following holds with probability $1 - o(1)$: for every fixed $\lambda \in \mathbb{R}$, if $p = \frac{1+\lambda n^{-1/3}}{d-1}$, then $L_1(G_{n,d,p}) = \Theta(n^{2/3})$. See also Riordan [23] for more precise results on $L_1(G_{n,d,p})$ in the critical window.

The case when d is an arbitrary function of n is much less understood. It follows from existing results in the literature[1] that for every $d \in \{3, \dots, n-1\}$, the critical probability for the existence of a linear order component in $G_{n,d,p}$ is $1/(d-1)$. Results inside the critical window for given d-regular graphs have also been obtained in the context of transitive graphs under the finite triangle condition [4] or under certain expansion conditions [18].

Finally, similar results have been obtained for irregular degree sequences whenever the average degree is bounded by a constant [3, 6, 7, 10].

In view that both the sparse regime (fixed $d \geq 3$) and the densest one ($d = n-1$) exhibit similar properties, Nachmias and Peres [20] suggested that the mean-field behaviour extends to

The first author was supported by the EPSRC, grant no. EP/M009408/1.

[1]The non-existence of a linear order component when $p \leq (1-\epsilon)p_{crit}$ follows from Proposition 1 in [20]. The existence of a linear order component when $p \geq (1+\epsilon)p_{crit}$ follows from the expansion properties of $G_{n,d}$ (see Corollary 2.8 in [13]) and the results on $(n, d, \lambda)-$graphs in [12].

every $d \in \{3, \dots, n-1\}$. In this paper we confirm this prediction in the critical window and thus answer the question posed by Benjamini for all $d \in \{3, \dots, n-1\}$.

Theorem 1. *Suppose $\lambda \in \mathbb{R}$ and $d, n \in \mathbb{N}$ such that $3 \le d \le n-1$ and n is sufficiently large. Let $p = \frac{1+\lambda n^{-1/3}}{d-1}$. Then for every sufficiently large $A = A(\lambda)$, we have*

$$\mathbb{P}[L_1(G_{n,d,p}) \notin [A^{-1}n^{2/3}, An^{2/3}]] \le 20A^{-1/2} .$$

The upper bound in Theorem 1 directly follows from the upper bound for d-regular graphs in Proposition 1 in [20]. The proof of the lower bound is more intricate and we devote the rest of the paper to it.

Most of the previous work on the component structure of $G_{n,d,p}$ uses the configuration model introduced by Bollobás in [2]. The configuration model, denoted by $G^*_{n,d}$, is a model of random d-regular multigraphs on n vertices. Conditional on $G^*_{n,d}$ being simple, one obtains the uniform distribution on $\mathcal{G}_{n,d}$. It is well-known (see for example [24]) that

$$\mathbb{P}[G^*_{n,d} \text{ simple}] = e^{-\Omega(d^2)} . \tag{1}$$

While $\mathbb{P}[G^*_{n,d} \text{ simple}]$ is constant for fixed $d \ge 3$, it quickly tends to 0 if d grows with n, and new ideas are needed to study $G_{n,d}$. A standard tool to estimate probabilities for $G_{n,d}$ when d grows with n is the switching method, introduced by McKay in [16]. For instance, this method has been used to estimate (1) for $d = o(\sqrt{n})$ [17] or to determine several combinatorial properties of $G_{n,d}$ when d grows with n [13].

The proof of the lower bound in Theorem 1 is based on the analysis of an exploration process in $G_{n,d,p}$ using the switching method. The central quantity that we track through the process is the number of edges between the explored and unexplored parts of the graph, denoted by X_t. Our proof relies on sharp estimations of the first and second moments of X_t.

This approach is inspired by recent developments of the switching method for the study of the component structure of random graphs with a given degree sequence [7, 11]. We take this opportunity to illustrate the use of our method with a simple proof that makes no assumptions on d.

The critical window. Theorem 1 shows that the critical window has width $\Omega(n^{-1/3})$. Proposition 1 in [20] implies that, as $\lambda \to -\infty$, the typical order of the largest component is $o(n^{2/3})$. Following analogous ideas as the ones used in the proof of Theorem 1, one obtains that, as $\lambda \to \infty$, the typical order of the largest component is $\omega(n^{2/3})$. More precisely, there exist constants $c, C > 0$ such that for every $3 \le d \le n-1$ and $\lambda > 0$, if $p = \frac{1+\lambda n^{-1/3}}{d-1}$, then

$$\mathbb{P}\left[L_1(G_{n,d,p}) \le c \cdot \lambda n^{2/3}\right] \le C\lambda^{-1} .$$

The proof of this statement is simpler than the proof of our main theorem, since the assumption $\lambda > 0$ implies that X_t has positive drift. In particular, the first part of the exploration process can be analysed using a first moment argument only and for the entire process it suffices to control the variance of X_t from above. It follows that the width of the critical window is $\Theta(n^{-1/3})$.

In its current form, our method does not give sharp estimates for $L_1(G_{n,d,p})$ in the barely subcritical and barely supercritical regimes. However, we believe that similar estimates as the ones in Lemma 6 hold in general and may be used to extend the results of Nachmias and Peres in [20] to all $d \in \{3, \dots, n-1\}$.

Diameter and Mixing Time. We present a consequence of Theorem 1. For a component $\mathcal{C}$, let $\mathrm{diam}(\mathcal{C})$ denote its diameter and let $T_{\mathrm{mix}}(\mathcal{C})$ denote the mixing time of the lazy random walk on $\mathcal{C}$. Theorem 1.2 in [19] implies the following corollary.

Corollary 2. *Suppose $\lambda \in \mathbb{R}$ and $d, n \in \mathbb{N}$ such that $3 \le d \le n-1$ and n is sufficiently large. Let $p = \frac{1+\lambda n^{-1/3}}{d-1}$. Let $\mathcal{C}$ be the largest component of $G_{n,d,p}$. Then, for every $\epsilon > 0$, there exists $A = A(\lambda, \epsilon)$ such that*

$$\mathbb{P}[\mathrm{diam}(\mathcal{C}) \notin [A^{-1}n^{1/3}, An^{1/3}]] < \epsilon .$$

and

$$\mathbb{P}[T_{\mathrm{mix}}(\mathcal{C}) \notin [A^{-1}n, An]] < \epsilon \,.$$

Organisation of the paper. The paper is organized as follows. In Section 2, we describe our exploration process of $G_{n,d,p}$ and introduce different quantities we will track during the process. In Section 3, we present our main combinatorial tool (switching method) and prove two technical lemmas. In Section 4, we use these lemmas to study a single step of the exploration process. Finally, in Section 5, we conclude with the proof of the lower bound in Theorem 1.

2. The exploration process

Before describing the exploration process, we briefly introduce some notation. For a graph G, a subset of vertices X of G, and a vertex u of G, we write $d_G(u)$ for the number of neighbours of u in G and $d_{G,X}(u)$ for the number of neighbours of u in G that belong to X. We also write $\Delta(G)$ for the maximum degree of G. Finally, for $p \in [0,1]$, we write G_p for the graph where each edge in G is independently retained with probability p.

We will use an exploration process to reveal the component structure of $G_{n,d,p}$. Let us denote the vertex set by V, which we equip with a linear order (from now on V is always a vertex set of size n). For technical reasons, we perform our exploration process not on $G_{n,d,p}$, but on what we call an input. An *input* is a tuple $(G, \mathfrak{S})$, where $G \in \mathcal{G}_{n,d}$ and $\mathfrak{S} = \{\sigma_v\}_{v \in V}$ is a collection of n permutations of length d. For each vertex of G, arbitrarily label the edges incident to it with distinct elements from $\{1, \dots, d\}$. Thus every edge receives two labels. In fact, we may think about this as a labelling of the semi-edges of G. Let $\mathcal{I}$ be the set of all inputs $(G, \mathfrak{S})$ where $G \in \mathcal{G}_{n,d}$ and $\mathfrak{S}$ is a collection of n permutations of length d. Observe that every graph in $G \in \mathcal{G}_{n,d}$ gives rise to exactly $(d!)^n$ inputs. Thus, choosing an input uniformly at random from $\mathcal{I}$ and ignoring the edge-labels is equivalent to choosing $G_{n,d}$. Let $\mathfrak{S}_{n,d}$ be a collection of n permutations of length d each chosen independently and uniformly at random. Hence, if an input is chosen uniformly at random from $\mathcal{I}$, then this input is distributed as $(G_{n,d}, \mathfrak{S}_{n,d})$.

Next, we describe our exploration process on an input $(G, \mathfrak{S})$. First, for every $uv \in E(G)$, we denote by $I(uv)$ the indicator random variable that is 1 if uv belongs to G_p (it percolates) and 0 otherwise. If $I(uv)$ is revealed, we say that the edge uv has been exposed. For each integer $t \geq 0$, the set S_t consists of the vertices explored up to time t (with $S_0 = \emptyset$); the bipartite graph F_t, with bipartition $(S_t, V \setminus S_t)$, consists of all edges in G between S_t and $V \setminus S_t$ that have been exposed and have failed to percolate; and the graph H_t, with vertex set S_t, consists of all edges in G within S_t, that is, $H_t := G[S_t]$. Let $\mathcal{H}_t$ be the history of all random choices we make until time t (which we will treat as an event).

We now describe how to obtain $\mathcal{H}_{t+1}$, given $\mathcal{H}_t$. Suppose there exists at least one vertex $u \in S_t$ such that $d_{H_t}(u) + d_{F_t}(u) < d$. Among all such vertices u, let v_{t+1} be the vertex which comes first in the linear order of V. Let w_{t+1} be the vertex $w \in V \setminus S_t$ with $v_{t+1}w \in E(G) \setminus E(F_t)$ that minimizes $\sigma_{v_{t+1}}(\ell(w))$, where $\ell(w)$ is the label of the semi-edge incident to v_{t+1} that corresponds to $v_{t+1}w$. Thereafter, we expose $v_{t+1}w_{t+1}$. If $I(v_{t+1}w_{t+1}) = 0$, then we set $S_{t+1} := S_t$, $Y_{t+1} := 0$, $Z_{t+1} := 0$ and we let F_{t+1} be the graph obtained from F_t by adding $v_{t+1}w_{t+1}$. If $I(v_{t+1}w_{t+1}) = 1$, then we set

$$S_{t+1} := S_t \cup \{w_{t+1}\}, \;\; Y_{t+1} := d_{F_t}(w_{t+1}), \;\; Z_{t+1} := d_{G,S_t}(w_{t+1}) - Y_{t+1} - 1,$$

and we let F_{t+1} be the graph obtained from F_t by deleting all edges incident to w_{t+1} and moving w_{t+1} to the other side of the bipartition. Since $H_{t+1} = G[S_{t+1}]$, we also reveal all the edges between w_{t+1} and S_t. Observe that Z_{t+1} counts the number of neighbours of w_{t+1} in $S_t \setminus \{v_{t+1}\}$ whose corresponding edge has not yet been exposed.

If $d_{H_t}(u) + d_{F_t}(u) = d$ for all $u \in S_t$, that is, every edge incident to a vertex in S_t has been exposed, then we pick a vertex $x \in V \setminus S_t$ that minimises $d_{F_t}(x)$ and set $w_{t+1} := x$, $S_{t+1} := S_t \cup \{w_{t+1}\}$, $Y_{t+1} := d_{F_t}(w_{t+1})$, $Z_{t+1} := 0$ and we let F_{t+1} be the graph obtained from F_t by deleting all edges incident to w_{t+1} and by moving w_{t+1} to the other side of the bipartition. Observe that, in any of the above-mentioned cases, $|E(F_{t+1})| \leq |E(F_t)| + 1$ and hence $|E(F_t)| \leq t$.

A crucial parameter of our exploration process is the number of edges between S_t and $V \setminus S_t$ which have not yet been exposed:

$$X_t := \sum_{u \in S_t} (d - d_{H_t}(u) - d_{F_t}(u)) \ .$$

For the sake of simplicity, we define $\eta_{t+1} := X_{t+1} - X_t$. If $X_t > 0$, then

$$\eta_{t+1} = -(1 - I(v_{t+1}w_{t+1})) + I(v_{t+1}w_{t+1})(d - 2 - Y_{t+1} - 2Z_{t+1}) \ , \tag{2}$$

and if $X_t = 0$, then

$$\eta_{t+1} = d - Y_{t+1} \ . \tag{3}$$

Note that Y_{t+1} and Z_{t+1} are measurable random variables given $\mathcal{H}_t$ and thus η_{t+1} is a predictable sequence with respect to $\mathcal{H}_t$.

3. The switching method and some applications

In this section we explain the switching method and we present two simple applications. In Lemma 3 we use the switching method to bound the probability from above that two vertices are adjacent. In Lemma 4 we provide an upper bound on the expectation of the number of neighbours of a vertex in a specified set of vertices.

Let G be a graph and let x_1, x_2, x_3, x_4 be distinct vertices of G. Suppose $x_1x_2, x_3x_4 \in E(G)$ and $x_1x_4, x_2x_3 \notin E(G)$. A *switching* on the 4-cycle $x_1x_2x_3x_4$ transforms G into a graph G' by deleting x_1x_2, x_3x_4 and adding x_1x_4, x_2x_3. Observe that the degree sequence of G is preserved by the switching. In particular, if G is d-regular, then so is G'. Moreover, the switching operation is reversible: if G can be transformed into G' by a switching, then G can be also obtained from G' by a switching on the same 4-cycle. Finally, there is a natural way to extend the notion of a switching from graphs to inputs by simply preserving the labels on each semi-edge.

Switchings can be used to obtain bounds on the probability that $G_{n,d}$ satisfies a certain property. Suppose $\mathcal{A}, \mathcal{B}$ are disjoint subsets of $\mathcal{G}_{n,d}$. Suppose that for every graph $G \in \mathcal{A}$, there are at least a switchings that transform G into a graph in $\mathcal{B}$ and for every graph $G' \in \mathcal{B}$, there are at most b switchings that transform G' into a graph in $\mathcal{A}$. By double-counting the number of switchings between $\mathcal{A}$ and $\mathcal{B}$, we obtain $a|\mathcal{A}| \leq b|\mathcal{B}|$. Thus $a\mathbb{P}[\mathcal{A}] \leq b\mathbb{P}[\mathcal{B}]$, where we define $\mathbb{P}[\mathcal{S}] := |\mathcal{S}|/|\mathcal{G}_{n,d}|$ for every $\mathcal{S} \subseteq \mathcal{G}_{n,d}$.

Lemma 3. *Suppose $d, n \in \mathbb{N}$ such that $3 \leq d \leq n/4$ and $S \subseteq V$ such that $|S| \leq n/6$. Let H be a graph with vertex set S and let F be a bipartite graph with vertex partition $(S, V \setminus S)$ with $\Delta(F \cup H) \leq d$. Let $u \in S$ and $v \in V \setminus S$ such that $uv \notin E(F)$. Then*

$$\mathbb{P}[uv \in E(G_{n,d}) \mid G_{n,d}[S] = H,\, F \subseteq G_{n,d}] \leq \frac{6(d - d_H(u) - d_F(u))}{n} \ .$$

Proof. Let $\mathcal{F}^+$ be the set of graphs $G \in \mathcal{G}_{n,d}$ such that $G[S] = H$, $F \subseteq G$ and $uv \in E(G)$, and let $\mathcal{F}^-$ be the set of graphs $G \in \mathcal{G}_{n,d}$ such that $G[S] = H$, $F \subseteq G$ but $uv \notin E(G)$. We will only perform switchings that involve edges and non-edges that are not contained in $E(H) \cup E(F)$. This ensures that the graph G' obtained from a switching also satisfies $G'[S] = H$ and $F \subseteq G'$.

Suppose $G \in \mathcal{F}^+$. In order to bound the number of switchings from below it suffices to switch on a cycle $uvxy$ that satisfies $xy \in E(G)$, $uy, vx \notin E(G)$, and $x, y \in V \setminus S$. There are at least $dn - 2d|S|$ ordered edges xy with both endpoints in $V \setminus S$. There are at most d^2 edges xy such that x is at distance at most 1 from v and at most d^2 edges xy such that y is at distance at most 1 from u. Thus, there are at least $dn - 2d|S| - 2d^2 \geq dn/6$ switchings that transform G into a graph in $\mathcal{F}^-$. Suppose now $G \in \mathcal{F}^-$. Then there are clearly at most $d \cdot (d - d_H(u) - d_F(u))$

switchings that transform G into a graph in $\mathcal{F}^+$. It follows that

$$\begin{aligned}\mathbb{P}[uv \in E(G_{n,d}) \mid G_{n,d}[S] = H,\, F \subseteq G_{n,d}] \\ &\leq \frac{d(d - d_H(u) - d_F(u))}{dn/6} \cdot \mathbb{P}[uv \notin E(G_{n,d}) \mid G_{n,d}[S] = H,\, F \subseteq G_{n,d}] \\ &\leq \frac{6(d - d_H(u) - d_F(u))}{n} .\end{aligned}$$

□

Lemma 4. *Suppose $d, n \in \mathbb{N}$ such that $3 \leq d \leq n/4$ and $S \subseteq V$ such that $|S| \leq n/6$. Let H be a graph with vertex set S and let F be a bipartite graph with vertex partition $(S, V \setminus S)$ with $\Delta(F \cup H) \leq d$. Let $v \in V \setminus S$. Then*

$$\mathbb{E}[d_{G,S}(v) - d_F(v) \mid G_{n,d}[S] = H,\ F \subseteq G_{n,d}] \leq 6d|S|/n.$$

Proof. For every $k \geq 0$, let $\mathcal{F}_k$ be the set of graphs $G \in \mathcal{G}_{n,d}$ such that $G[S] = H$, $F \subseteq G$, and $d_{G,S}(v) - d_F(v) = k$. As in Lemma 3, we will only perform switchings using edges and non-edges that are not contained in $E(H) \cup E(F)$.

Consider a graph in $\mathcal{F}_k$. There are at most $(d - d_F(v)) \cdot d|S| \leq d^2|S|$ switchings that lead to a graph in $\mathcal{F}_{k+1}$. For every graph in $\mathcal{F}_{k+1}$, we can use a switching on a cycle $uvxy$ that satisfies $uv, xy \in E(G) \setminus E(F)$, $uy, vx \notin E(G)$ and $u \in S$, and $v, x, y \in V \setminus S$. There are $k + 1$ choices for uv and, for any particular choice of uv, there are at least $dn - 2d|S| - 2d^2 \geq dn/6$ choices for the (ordered) edge xy. Hence, there are at least $(k + 1)dn/6$ switchings that lead to a graph in $\mathcal{F}_k$. Thus, for every $k \geq 0$, we obtain

$$\mathbb{P}[\mathcal{F}_{k+1}] \leq \frac{6d|S|/n}{(k+1)} \cdot \mathbb{P}[\mathcal{F}_k] . \tag{4}$$

Let X be a Poisson distributed random variable with mean $6d|S|/n$. Lemma 3.4 in [15] together with (4) implies that for every $m \geq 0$

$$\mathbb{P}[d_{G,S}(v) - d_F(v) \geq m \mid G_{n,d}[S] = H,\ F \subseteq G_{n,d}] \leq \mathbb{P}[X \geq m] ,$$

which implies the statement of the lemma. □

4. Analysis of the exploration process

In this section we show how to control the expectation of η_t and η_t^2. We first use Lemmas 3 and 4 to bound the expectation of Y_{t+1} and Z_{t+1} from above.

Lemma 5. *Suppose $d, n \in \mathbb{N}$ such that $3 \leq d \leq n - 1$ and n is sufficiently large. Fix $p \in [0, 1]$. Consider the exploration process described above on $(G_{n,d}, \mathfrak{S}_{n,d})$ with percolation probability p and suppose $t \leq dn^{2/3}$. Conditional on $\mathcal{H}_t$ satisfying $|S_t| \leq 5n^{2/3}$, we have*

$$\mathbb{E}[Y_{t+1}|\mathcal{H}_t] \leq 20dn^{-1/3} \quad \textit{and} \quad \mathbb{E}[Z_{t+1}|\mathcal{H}_t] \leq 180dn^{-1/3} .$$

Proof. If $\mathcal{H}_t$ satisfies $X_t = 0$, then $Y_{t+1} \leq t/(n - |S_t|) \leq 2dn^{-1/3}$ by our choice of w_{t+1} (we always choose the vertex x that minimises $d_{F_t}(x)$) and $|E(F_t)| \leq t$. Note that $Z_{t+1} = 0$ by definition. Hence we may assume from now on that $X_t > 0$.

Note that if $d \geq n/4$, then the lemma follows directly from the fact that $Y_{t+1} \leq |S_t| \leq 5n^{2/3} \leq 20dn^{-1/3}$, and similarly for Z_{t+1}. Thus, in the following we assume that $d \leq n/4$.

Given $w \in V \setminus S_t$ such that $v_{t+1}w \notin E(F_t)$, we apply Lemma 3 with $S = S_t$, $F = F_t$, $H = H_t$, $u = v_{t+1}$ and $v = w$ to obtain

$$\mathbb{P}[v_{t+1}w \in E(G_{n,d}) \mid v_{t+1}w \notin E(F_t), \mathcal{H}_t] \leq \frac{6(d - d_{H_t}(v_{t+1}) - d_{F_t}(v_{t+1}))}{n} .$$

Observe that we run our exploration process on inputs. In order to apply Lemma 3, we fix the semi-edge labelings and perform switchings on the graphs.

Since $\sigma_{v_{t+1}}$ is a random permutation, each edge incident to v_{t+1} that is not contained in $E(F_t) \cup E(H_t)$ is chosen with the same probability to continue the exploration process. Hence, given that $v_{t+1}w \in E(G_{n,d}) \setminus E(F_t)$, the probability that $w_{t+1} = w$ is precisely $(d - d_{H_t}(v_{t+1}) - d_{F_t}(v_{t+1}))^{-1}$. Therefore,

$$\begin{aligned}&\mathbb{P}[w_{t+1} = w \mid v_{t+1}w \notin E(F_t), \mathcal{H}_t]\\ &= \mathbb{P}[w_{t+1} = w \mid v_{t+1}w \in E(G_{n,d}) \setminus E(F_t), \mathcal{H}_t] \cdot \mathbb{P}[v_{t+1}w \in E(G_{n,d}) \mid v_{t+1}w \notin E(F_t), \mathcal{H}_t] \leq \frac{6}{n} .\end{aligned}$$

Since $\mathbb{P}[w_{t+1} = w \mid v_{t+1}w \in E(F_t), \mathcal{H}_t] = 0$, it follows that for every $w \in V \setminus S_t$

$$\mathbb{P}[w_{t+1} = w \mid \mathcal{H}_t] \leq \frac{6}{n} . \tag{5}$$

Using that $|E(F_t)| \leq t$, we conclude

$$\mathbb{E}[Y_{t+1}|\mathcal{H}_t] = \sum_{w \in V \setminus S_t} d_{F_t}(w)\mathbb{P}[w_{t+1} = w|\mathcal{H}_t] \overset{(5)}{\leq} \frac{6}{n} \sum_{w \in V \setminus S_t} d_{F_t}(w) \leq \frac{6}{n} \cdot t \leq 6dn^{-1/3} .$$

We now prove the second statement. Given $w \in V \setminus S_t$ with $\mathbb{P}[w_{t+1} = w \mid \mathcal{H}_t] > 0$ (that is, $v_{t+1}w \notin E(F_t)$), we apply Lemma 4 with $S = S_t$, F obtained from F_t by adding $v_{t+1}w$, $H = H_t$, and $v = w$, to obtain

$$\begin{aligned}\mathbb{E}[Z_{t+1}|\mathcal{H}_t] &= \sum_{w \in V \setminus S_t} \mathbb{E}[Z_{t+1}|w_{t+1} = w, v_{t+1}w \notin E(F_t), \mathcal{H}_t]\mathbb{P}[w_{t+1} = w \mid v_{t+1}w \notin E(F_t), \mathcal{H}_t]\\ &\overset{(5)}{\leq} \sum_{w \in V \setminus S_t} \frac{6d|S_t|}{n} \cdot \frac{6}{n} \leq 180dn^{-1/3} .\end{aligned}$$

□

Lemma 6. *Suppose $\mu \geq 0$ and $d, n \in \mathbb{N}$ such that $3 \leq d \leq n - 1$ and n is sufficiently large. Consider the exploration process described above on $(G_{n,d}, \mathfrak{S}_{n,d})$ with $p = \frac{1-\mu n^{-1/3}}{d-1}$ and suppose $t \leq dn^{2/3}$. Conditional on $|S_t| \leq 5n^{2/3}$, then*

$$\mathbb{E}[\eta_{t+1}|\mathcal{H}_t] \geq -(570 + \mu)n^{-1/3} \quad \text{and} \quad \mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] \geq d/4 .$$

Moreover, if $X_t > 0$, then $\mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] \leq d$.

Proof. First assume that $X_t > 0$. Recall that for any $\mathcal{H}_t$ and for any edge uv that has not been exposed yet, we have $\mathbb{E}[I(uv) \mid \mathcal{H}_t] = p = (1 - \mu n^{-1/3})/(d - 1)$. Recall that Y_{t+1} and Z_{t+1} are measurable with respect to $\mathcal{H}_t$. Taking conditional expectations on (2) and using Lemma 5, we obtain

$$\begin{aligned}\mathbb{E}[\eta_{t+1}|\mathcal{H}_t] &= -\left(1 - \frac{1 - \mu n^{-1/3}}{d - 1}\right) + \frac{1 - \mu n^{-1/3}}{d - 1}(d - 2 - \mathbb{E}[Y_{t+1}|\mathcal{H}_t] - 2\mathbb{E}[Z_{t+1}|\mathcal{H}_t])\\ &\geq -\frac{\mathbb{E}[Y_{t+1}|\mathcal{H}_t] + 2\mathbb{E}[Z_{t+1}|\mathcal{H}_t]}{d - 1} - \mu n^{-1/3}\\ &\geq -\frac{380dn^{-1/3}}{d - 1} - \mu n^{-1/3} \geq -(570 + \mu)n^{-1/3} ,\end{aligned}$$

since $d \geq 3$.

Again, by Lemma 5 and (2), we obtain

$$\begin{aligned}\mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] &= \left(1 - \frac{1 - \mu n^{-1/3}}{d - 1}\right)(-1)^2 + \frac{1 - \mu n^{-1/3}}{d - 1}\mathbb{E}[(d - 2 - Y_{t+1} - 2Z_{t+1})^2 \mid \mathcal{H}_t]\\ &\geq \frac{d - 2}{d - 1} + \frac{(1 - \mu n^{-1/3})(d - 2)^2}{d - 1} - \frac{2(d - 2)(\mathbb{E}[Y_{t+1}|\mathcal{H}_t] + 2\mathbb{E}[Z_{t+1}|\mathcal{H}_t])}{d - 1}\\ &\geq (1 - \mu n^{-1/3})(d - 2) - 2(\mathbb{E}[Y_{t+1}|\mathcal{H}_t] + 2\mathbb{E}[Z_{t+1}|\mathcal{H}_t])\\ &\geq (1 - \mu n^{-1/3})(d - 2) - 760dn^{-1/3}\\ &\geq d/4 ,\end{aligned}$$

where the last inequality holds since $d \geq 3$ and n is sufficiently large. Observe that $\mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] \leq d$ follows from a similar argument as $(d-2-Y_{t+1}-2Z_{t+1})^2 \leq (d-2)^2$.

If $X_t = 0$, then clearly $\mathbb{E}[\eta_{t+1}|\mathcal{H}_t] \geq 0$ and, since $\mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] = \mathbb{E}[(d-Y_{t+1})^2|\mathcal{H}_t]$, similarly as before, one can prove that $\mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] \geq d/4$. □

Lemma 7. *Suppose $\mu \geq 0$ and $d, n \in \mathbb{N}$ such that $3 \leq d \leq n-1$ and n is sufficiently large. Consider the exploration process described above on $(G_{n,d}, \mathfrak{S}_{n,d})$ with $p = \frac{1-\mu n^{-1/3}}{d-1}$. Then, for every fixed $\delta > 0$ and all $0 \leq t_1 \leq t_2 \leq 5dn^{2/3}$, we have*

$$\mathbb{P}\left[|S_{t_2} \setminus S_{t_1}| - \frac{t_2 - t_1}{d-1} \geq -\delta n^{2/3}\right] = 1 - o(n^{-2}) \quad \text{and}$$
$$\mathbb{P}\left[|S_{t_2} \setminus S_{t_1}| - \frac{t_2 - t_1}{d-1} - \left\lceil \frac{t_2}{5d/6} \right\rceil \leq \delta n^{2/3}\right] = 1 - o(n^{-2}) \ .$$

Proof. We add a vertex to S_t either if $I(v_{t+1}w_{t+1}) = 1$ or if we start exploring a new component of $G_{n,d,p}$ at time $t+1$. Thus, $|S_{t_2} \setminus S_{t_1}|$ stochastically dominates a binomial random variable with parameters $t_2 - t_1$ and $(1-\mu n^{-1/3})/(d-1)$. A standard application of Chernoff's inequality implies the first statement.

Let $A_t \subseteq S_t$ be the set of vertices that start a new component in $G_{n,d,p}$. For every $0 \leq t \leq 5dn^{2/3}$, let $a_t := |A_t|$, let $c_t := |S_t \setminus A_t|$ and let $b_t := |S_t \setminus (S_{t_1} \cup A_t)|$. Observe that c_t is stochastically dominated by a binomial random variable with parameters t and $1/(d-1)$. Using Chernoff's inequality, we have $c_t \leq 8n^{2/3}$ with probability $1 - o(n^{-2})$ for any $t \leq 5dn^{2/3}$.

We claim that for every $0 \leq t \leq 5dn^{2/3}$ and conditional on $c_t \leq 8n^{2/3}$, we have $a_t \leq \lceil \frac{t}{5d/6} \rceil$. Indeed, the claim is true for $t \in \{0, 1\}$. Assume that $t \geq 2$ and that the claim holds for every $t' \in \{0, \dots, t-1\}$. If $X_{t-1} > 0$, then $a_t = a_{t-1}$ and we are done. Thus, assume that $X_{t-1} = 0$. Let s be the largest integer $s' \in \{0, \dots, t-2\}$ such that $X_{s'} = 0$ (it exists since $X_0 = 0$ and $t \geq 2$). Recall that w_{s+1} is a vertex $x \in V \setminus S_s$ that minimises $d_{F_s}(x)$. It follows that

$$d_{F_s}(w_{s+1}) \leq \frac{|E(F_s)|}{n-(a_s+c_s)} \leq \frac{s}{n - \lceil s/(5d/6) \rceil - 8n^{2/3}} \leq \frac{d}{6} \ ,$$

provided that n is large enough. Hence, $X_{s+1} \geq 5d/6$ and the process will not start a new component for the next $5d/6$ steps. In particular, $s + 5d/6 \leq t$. This implies $a_t = a_s + 1 \leq \lceil \frac{s}{5d/6} \rceil + 1 \leq \lceil \frac{t}{5d/6} \rceil$.

Since $|S_{t_2} \setminus S_{t_1}| \leq a_{t_2} + b_{t_2}$, the second part of the lemma now follows from the upper bound on a_{t_2} (which holds as we assume $c_t \leq 8n^{2/3}$) and an upper bound on b_{t_2} obtained by Chernoff's inequality. □

5. Proof of Theorem 1

As we mentioned in the introduction, due to the result of Nachmias and Peres, we only need to prove a lower bound. Since it suffices to prove the lower bound of the statement for $\lambda \leq 0$, we use the definition $\mu := -\lambda$. We now present a brief overview of the proof. In the first phase, we show that with probability at least $1 - A^{-1/2}$, the process X_t exceeds $A^{-1/4}dn^{1/3}$ in the first $dn^{2/3}$ steps. In the second phase and conditional on the success of the first phase, we show that X_t stays positive for at least $2A^{-1}dn^{2/3}$ steps with probability at least $1 - A^{-1/2}$. From standard concentration inequalities, this gives the existence of a component of order at least $A^{-1}n^{2/3}$, concluding the proof. This proof strategy was introduced by Nachmias and Peres to prove the same statement for fixed $d \geq 3$ [20] and for $d = n-1$ [21]. We remark that, in comparison to [20], our analysis of the exploration process is simpler, as we do not need to track the number of vertices $x \in V \setminus S_t$ which satisfy $d_{F_t}(x) = k$ for $k \in \{0, 1, \dots, d\}$. If $d \geq 3$ is fixed, as in [20], almost every vertex x satisfies $d_{F_t}(x) \in \{0, 1\}$. However, this is no longer true if d is an arbitrary function of n. We avoid the technicalities involved with this issue by averaging over the values of $d_{F_t}(x)$.

First phase: We start with the definition of a few parameters. Let $h := A^{-1/4}dn^{1/3}$, $T_1 := 5dn^{2/3}/6$ and $T_2 := 2A^{-1}dn^{2/3}$. In addition, we define the following stopping times:

$$\begin{aligned}\tau_h &:= \min\{t : X_t \geq h\} \wedge T_1\\ \tau_S^1 &:= \min\{t : |S_t| \geq 3n^{2/3}\}\\ \tau_1 &:= \tau_h \wedge \tau_S^1 .\end{aligned}$$

Recall that $X_{t+1} = \eta_{t+1} + X_t$. Note also that for every $t < \tau_1$, we have $X_t \leq h$ and $|S_t| \leq 5n^{2/3}$. Hence, Lemma 6 implies that

$$\mathbb{E}[X_{t+1}^2 - X_t^2|\mathcal{H}_t] \geq \mathbb{E}[\eta_{t+1}^2|\mathcal{H}_t] + 2\mathbb{E}[\eta_{t+1}X_t|\mathcal{H}_t] \geq d/4 - 2\cdot(570+\mu)n^{-1/3}h \geq d/5 ,$$

provided that A is large enough with respect to μ (and thus, with respect to λ). Hence $X_{t\wedge\tau_1}^2 - (t\wedge\tau_1)d/5$ is a submartingale. By the Optional Stopping theorem for submartingales (see for example [9] p.491), $\mathbb{E}[X_{\tau_1}^2 - \frac{d}{5}\tau_1] \geq \mathbb{E}[X_0^2] = 0$, which implies that $\mathbb{E}[\tau_1] \leq \frac{5}{d}\mathbb{E}[X_{\tau_1}^2]$. Since $X_{\tau_1}^2 \leq (h+d)^2 \leq 2h^2$, we obtain

$$\mathbb{P}[\tau_1 = T_1] \leq \frac{\mathbb{E}[\tau_1]}{T_1} \leq \frac{5\mathbb{E}[X_{\tau_1}^2]}{dT_1} \leq \frac{10h^2}{dT_1} = 12A^{-1/2} .$$

By Lemma 7 with $t_1 = 0$ and $t_2 = T_1$, we have $\mathbb{P}[\tau_S^1 \leq T_1] = o(1)$. Thus

$$\mathbb{P}[\{\tau_h = T_1\} \cup \{\tau_S^1 \leq \tau_h\}] \leq \mathbb{P}[\tau_1 = T_1] + \mathbb{P}[\tau_S^1 \leq T_1] \leq 12A^{-1/2} + o(1) \leq 13A^{-1/2} . \quad (6)$$

We conclude that the event $\mathcal{E} := \{\tau_h < T_1, \tau_h < \tau_S^1\}$ holds with probability at least $1 - 13A^{-1/2}$. In particular, with probability at least $1 - 13A^{-1/2}$, the random process X_t exceeds h before time T_1.

Second phase: Write $\mathbb{P}_*$ and $\mathbb{E}_*$ for the probability and the expectation conditional on $\mathcal{E}$. We define

$$\begin{aligned}\tau_0 : &= \min\{t : X_{\tau_h+t} = 0\} \wedge T_2\\ \tau_S^2 : &= \min\{t : |S_{\tau_h+t} \setminus S_{\tau_h}| \geq 2n^{2/3}\}\\ \tau_2 : &= \tau_0 \wedge \tau_S^2 .\end{aligned}$$

Consider the random variable

$$W_t := h - \min\{h, X_{\tau_h+t}\} .$$

Hence

$$\begin{aligned}W_{t+1}^2 - W_t^2 &\leq (h - \min\{h, X_{\tau_h+t}\} - \eta_{\tau_h+t+1})^2 - (h - \min\{h, X_{\tau_h+t}\})^2\\ &= \eta_{\tau_h+t+1}^2 - 2\eta_{\tau_h+t+1}(h - \min\{h, X_{\tau_h+t}\})\\ &\leq \eta_{\tau_h+t+1}^2 - 2\eta_{\tau_h+t+1}h .\end{aligned}$$

If $t < \tau_2$ and n is sufficiently large, we can apply Lemma 6 and this leads to (provided A is sufficiently large with respect to μ)

$$\mathbb{E}_*[W_{t+1}^2 - W_t^2 \mid \mathcal{H}_{\tau_h+t}] \leq d + 2\cdot(570+\mu)n^{-1/3}\cdot h \leq 2d .$$

Thus, $W_{t\wedge\tau_2}^2 - 2d(t\wedge\tau_2)$ is a supermartingale. Similar as before, we use the Optimal Stopping theorem to conclude that

$$\mathbb{E}_*[W_{\tau_2}^2] \leq 2d\mathbb{E}_*[\tau_2] \leq 2dT_2 .$$

Thus

$$\begin{aligned}\mathbb{P}_*[\tau_2 < T_2] &= \mathbb{P}_*[\tau_0 < T_2, \tau_S^2 > T_2] + \mathbb{P}_*[\tau_S^2 \leq T_2]\\ &\leq \mathbb{P}_*[W_{\tau_2} \geq h] + \mathbb{P}_*[|S_{\tau_h+T_2} \setminus S_{\tau_h}| \geq 2n^{2/3}]\\ &\leq \mathbb{P}_*[W_{\tau_2}^2 \geq h^2] + o(1)\\ &\leq \frac{\mathbb{E}_*[W_{\tau_2}^2]}{h^2} + o(1) \leq 5A^{-1/2} ,\end{aligned}$$

where we used Lemma 7 with $t_1 = \tau_h$ and $t_2 = \tau_h + T_2$ for the second inequality. (Observe that we cannot apply Lemma 7 directly, because we assume $\mathcal{E}$ holds and τ_h is a random time. However, as $\tau_h \le T_1$, a simple union bound with $t_1 = k$ and $t_2 = k + T_2$ for all $k \le T_1$ together with the fact that $\mathbb{P}[\mathcal{E}] \ge 1 - 13A^{-1/2} \ge 1/2$, yields the desired result.) It follows that

$$\begin{aligned}\mathbb{P}[\{\tau_2 < T_2\} \cup \{\tau_h = T_1\} \cup \{\tau_S^1 \le \tau_h\}] &\le \mathbb{P}[\{\tau_h = T_1\} \cup \{\tau_S^1 \le \tau_h\}] + \mathbb{P}_*[\tau_2 < T_2] \\ &\overset{(6)}{\le} 13A^{-1/2} + 5A^{-1/2} = 18A^{-1/2}\,.\end{aligned}$$

Since all the vertices explored from time τ_h to $\tau_h + \tau_2$ belong to the same component of $G_{n,d,p}$, there exists a component of size at least $|S_{\tau_h+\tau_2} \setminus S_{\tau_h}|$. As $\tau_2 = T_2 = 2A^{-1}dn^{2/3}$ with probability at least $1 - 18A^{-1/2}$, by Lemma 7 with $t_1 = \tau_h$ and $t_2 = \tau_h + T_2$ (as above, strictly speaking, we apply Lemma 7 with $t_1 = k$ and $t_2 = k + T_2$ for all $k \le T_1$ and use the fact that $\mathbb{P}[\mathcal{E}] \ge 1/2$) with probability at least $1 - 18A^{-1/2} - o(1) \ge 1 - 19A^{-1/2}$, there exists a component of size at least $A^{-1}n^{2/3}$.

Acknowledgements: The authors want to thank Nikolaos Fountoulakis, Michael Krivelevich, and Asaf Nachmias for fruitful discussions on the topic and the anonymous referees for their valuable comments.

References

1. N. Alon, I. Benjamini, and A. Stacey, *Percolation on finite graphs and isoperimetric inequalities*, Ann. Probab. (2004), 1727–1745.
2. B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, Europ. J. Combin. **1** (1980), 311–316.
3. B. Bollobás and O. Riordan, *An old approach to the giant component problem*, J. Combin. Theory (Series B) **113** (2015), 236–260.
4. C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade, and J. Spencer, *Random subgraphs of finite graphs. I. The scaling window under the triangle condition*, Random Structures Algorithms **27** (2005), 137–184.
5. P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61.
6. N. Fountoulakis, *Percolation on sparse random graphs with given degree sequence*, Internet Math. **4** (2007), 329–356.
7. N. Fountoulakis, F. Joos, and G. Perarnau, *Percolation on random graphs with a fixed degree sequence*, arXiv:1611.08496 (2016).
8. A. Goerdt, *The giant component threshold for random regular graphs with edge faults*, Theoret. Comput. Sci. **259** (2001), 307–321.
9. G. Grimmett and D. Stirzaker, *Probability and random processes*, Oxford University Press, 2001.
10. S. Janson, *On percolation in random graphs with given vertex degrees*, Electron. J. Probab. **14** (2009), 87–118.
11. F. Joos, G. Perarnau, D. Rautenbach, and B. Reed, *How to determine if a random graph with a fixed degree sequence has a giant component*, to appear in Probability Theory and Related Fields (2017).
12. M. Krivelevich and B. Sudakov, *The phase transition in random graphs: A simple proof*, Random Structures Algorithms **43** (2013), 131–138.
13. M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald, *Random regular graphs of high degree*, Random Structures Algorithms **18** (2001), 346–363.
14. T. Łuczak, B. Pittel, and J. C. Wierman, *The structure of a random graph at the point of the phase transition*, Trans. Amer. Math. Soc. **341** (1994), 721–748.
15. C. McDiarmid, *Connectivity for random graphs from a weighted bridge-addable class*, Electron. J. Combin. **19** (2012), no. 4, Paper 53, 20.
16. B. D. McKay, *Asymptotics for symmetric 0-1 matrices with prescribed row sums*, Ars Combin **19** (1985), 15–25.
17. B. D. McKay and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$*, Combinatorica **11** (1991), 369–382.
18. A. Nachmias, *Mean-field conditions for percolation on finite graphs*, Geom. Funct. Anal. **19** (2009), 1171–1194.
19. A. Nachmias and Y. Peres, *Critical random graphs: diameter and mixing time*, Ann. Probab. (2008), 1267–1286.
20. ———, *Critical percolation on random regular graphs*, Random Structures Algorithms **36** (2010), 111–148.
21. ———, *The critical random graph, with martingales*, Israel J. Math. **176** (2010), 29–41.
22. B. Pittel, *Edge percolation on a random regular graph of low degree*, Ann. Probab. **36** (2008), 1359–1389.
23. O. Riordan, *The phase transition in the configuration model*, Comb. Probab. Comput. **21** (2012), 265–299.
24. N. C. Wormald, *Models of random regular graphs*, Surveys in combinatorics, 1999 (Canterbury), London Math. Soc. Lecture Note Ser., vol. 267, Cambridge Univ. Press, Cambridge, 1999, pp. 239–298.

Version November 12, 2017

Felix Joos
<f.joos@bham.ac.uk>
Guillem Perarnau
<g.perarnau@bham.ac.uk>
School of Mathematics, University of Birmingham, Birmingham
United Kingdom