

# Large matchings and nearly spanning, nearly regular subgraphs of random subgraphs

Sahar Diskin <sup>\*</sup>      Joshua Erde <sup>†</sup>      Mihyun Kang <sup>†</sup>  
Michael Krivelevich <sup>\*</sup>

## Abstract

Given a graph  $G$  and  $p \in [0, 1]$ , the random subgraph  $G_p$  is obtained by retaining each edge of  $G$  independently with probability  $p$ . We show that for every  $\epsilon > 0$ , there exists a constant  $C > 0$  such that the following holds. Let  $d \geq C$  be an integer, let  $G$  be a  $d$ -regular graph and let  $p \geq \frac{C}{d}$ . Then, with probability tending to one as  $|V(G)|$  tends to infinity, there exists a matching in  $G_p$  covering at least  $(1 - \epsilon)|V(G)|$  vertices.

We further show that for a wide family of  $d$ -regular graphs  $G$ , which includes the  $d$ -dimensional hypercube, for any  $p \geq \frac{\log^5 d}{d}$  with probability tending to one as  $d$  tends to infinity,  $G_p$  contains an induced subgraph on at least  $(1 - o(1))|V(G)|$  vertices, whose degrees are tightly concentrated around the expected average degree  $dp$ .

## 1 Introduction

A classical result of Erdős and Rényi [12] states that  $p = \frac{\log n}{n}$  is the threshold for the existence of a *perfect* matching (that is, a matching covering all but at most one vertex) in  $G(n, p)$ <sup>1</sup>, which also coincides with the connectivity threshold (see also [5] for a hitting time result). Below this threshold, it is not hard to show that with probability tending to one as  $n$  tends to infinity a fixed proportion of the vertices are isolated and will not be covered in any matching. On the other hand, it follows from a celebrated result of Karp and Sipser [14] from 1981 that, when  $p = \frac{C}{n}$  for a large enough constant  $C$ ,  $G(n, p)$  contains a matching on  $(1 - o_C(1))n$  vertices with probability tending to one as  $n$  tends to infinity. Subsequent work by Frieze [13] gave a precise estimate of the asymptotic proportion of vertices which are not covered by a largest matching in this regime.

The binomial random graph  $G(n, p)$  is an instance of the model of *bond percolation*. Given a host graph  $G$  and a probability  $p \in [0, 1]$ , we form the random subgraph  $G_p \subseteq G$  by retaining each edge of  $G$  independently with probability  $p$  (indeed,  $G(n, p)$  is equivalent to performing bond percolation on the complete graph  $K_n$ ). In a qualitative sense, our first main result extends the result of Karp and Sipser [14] to any  $d$ -regular graph.

**Theorem 1.** *For every  $\epsilon > 0$ , there exists a constant  $C > 0$  such that the following holds. Let  $d \in \mathbb{N}, d \geq C$ , let  $G$  be a  $d$ -regular graph on  $n$  vertices, and let  $p \geq \frac{C}{d}$ . Then, with probability tending to one as  $n$  tends to infinity, there exists a matching in  $G_p$  covering at least  $(1 - \epsilon)n$  vertices.*

In addition to the typical existence of large matchings in percolated  $d$ -regular graphs, we also explore typical structural properties of the hypercube under percolation.

In the setting of  $G(n, p)$ , finding a large  $k$ -regular subgraph when  $p = \frac{C}{n}$  has been extensively studied (see, for example, [15, 17] and references therein). A natural variant is to try and find

<sup>\*</sup>School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel. Emails: sahardiskin@mail.tau.ac.il, krivelev@tauex.tau.ac.il.

<sup>†</sup>Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria. Emails: erde@math.tugraz.at, kang@math.tugraz.at.

<sup>1</sup>In fact, Erdős and Rényi worked in the closely related *uniform* random graph model  $G(n, m)$ .

a *nearly regular*, nearly spanning subgraph (see [1] for an extremal variant of this question). We consider this question in the setting of random subgraphs of the hypercube. Recall that the  $d$ -dimensional binary hypercube is the graph whose vertex set is  $\{0,1\}^d$ , and where two vertices are connected if and only if their Hamming distance is one. In our second main result, we establish the typical existence of a nearly regular, nearly spanning induced subgraph of  $Q_p^d$ , whose degrees are tightly concentrated around the expected degree  $dp$ .

**Theorem 2.** *For every  $\epsilon > 0$ , there exists  $d_0 \in \mathbb{N}$  such that the following holds for every integer  $d \geq d_0$ . Let  $p \geq \frac{\log^5 d}{d}$ . Then, with probability tending to one as  $d$  tends to infinity there exists an induced subgraph  $H \subseteq Q_p^d$ , such that  $|V(H)| \geq (1 - \epsilon)2^d$  and for every  $v \in V(H)$ ,  $\left|1 - \frac{d_H(v)}{dp}\right| \leq \epsilon$ .*

We note that the proofs of the theorems are quite short. Further, let us remark that we have not tried to optimise the polylogarithmic dependence of  $p$  on  $d$  in Theorem 2. Finally, we note that we present the proof for  $Q_p^d$ , but, as will be expanded upon in the discussion section, Theorem 2 holds for a fairly wide family of graphs (see Theorem 7).

In Section 2 we prove Theorems 1 and 2. In Section 3 we discuss our results and avenues for future research.

## 2 Proofs of Theorems 1 and 2

Given a graph  $G$  and  $v \in V(G)$ , we denote by  $d_G(v)$  the degree of  $v$  in  $G$ . For  $k \in \mathbb{N}$ , we denote by  $N_G^k(v)$  the set of vertices at distance exactly  $k$  from  $v$  in  $G$ , where  $N_G(v) := N_G^1(v)$ . When the graph  $G$  is clear from the context, we will omit the subscripts. Further, given  $S \subseteq V(G)$ , we denote by  $G[S]$  the subgraph of  $G$  induced by the vertices of  $S$ . Given  $x, y, z \in \mathbb{R}$  we write  $x = y \pm z$  as shorthand for  $x \in [y - z, y + z]$ . Throughout the paper, all logarithms are in the natural base.

We will make use of a typical Chernoff-type tail bound on the binomial distribution (see, for example, Appendix A in [2]).

**Lemma 3.** *Let  $d \in \mathbb{N}$ , let  $p \in [0, 1]$ , and let  $X \sim \text{Bin}(d, p)$ . Then for any  $0 < t \leq \frac{dp}{2}$ ,*

$$\mathbb{P}[X \neq dp \pm t] \leq 2 \exp \left\{ -\frac{t^2}{3dp} \right\}.$$

We will also utilise a variant of the well-known Azuma-Hoeffding inequality (see, for example, Chapter 7 in [2]),

**Lemma 4.** *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  $X_i$  distributed according to independent Bernoulli( $p$ ). Let  $f : \Lambda \rightarrow \mathbb{R}$  be a function such that there exists  $K \in \mathbb{R}$  such that for every  $x, x' \in \Lambda$  which differ only in one coordinate,  $|f(x) - f(x')| \leq K$ . Then, for every  $t \geq 0$ ,*

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

### 2.1 Proof of Theorem 1

Fix  $\epsilon > 0$ . Let  $\delta := \delta(\epsilon) > 0$  and  $C := C(\delta, \epsilon) = C(\epsilon) > 0$  be constants satisfying that  $C \exp \left\{ -\frac{\delta^2 C}{16} \right\} \leq \frac{\epsilon}{4}$  and  $\frac{C - \epsilon}{(1 + \delta)C} \geq 1 - \epsilon$ . Note that by monotonicity, we may assume  $p = \frac{C}{d}$ . Let  $V_0$  be the set of vertices of degree at least  $(1 + \delta)C$ , and let  $E_0$  be the set of edges touching  $V_0$ , that is,

$$V_0 := \{v \in V(G) : d_{G_p}(v) \geq (1 + \delta)C\} \quad \text{and} \quad E_0 := \{uv \in E(G_p) : \{u, v\} \cap V_0 \neq \emptyset\}.$$

For any  $e \in E(G)$  the probability that  $e \in E(G_p)$  and at least one of its endpoints is in  $V_0$  is at most  $p \cdot 2\mathbb{P}\left(\text{Bin}\left(d, \frac{C}{d}\right) \geq (1 + \delta)C - 1\right)$ , which is bounded above by  $\frac{C}{d} \cdot 4 \exp\left\{-\frac{\delta^2 C}{4}\right\}$  by Lemma 3 (where we assumed that  $\delta C \geq 9$ ). Hence, using  $|E(G)| = \frac{dn}{2}$ , we have  $\mathbb{E}[|E_0|] \leq 2C \exp\left\{-\frac{\delta^2 C}{4}\right\} n \leq \frac{\varepsilon}{8}n$ . Now, note that adding/removing any edge can change  $|V_0|$  by at most two, and thus  $|E_0|$  by at most  $2(1 + \delta)C$ . Hence, by Lemma 4 (with  $t = \frac{\varepsilon n}{8}$ ,  $K = 2(1 + \delta)C$ , and  $m = \frac{nd}{2}$ ), and using  $\exp\left\{-\frac{\delta^2 C}{4}\right\} \leq \frac{\varepsilon}{16C}$  we obtain

$$\mathbb{P}\left[|E_0| \geq \frac{\varepsilon}{4}n\right] \leq 2 \exp\left\{-\frac{(\varepsilon n/8)^2}{2 \cdot 4(1 + \delta)^2 C^2 \cdot nd/2 \cdot C/d}\right\} = \exp\{-\Omega(n)\} = o_n(1).$$

By similar (and simpler) arguments, with probability  $1 - o_n(1)$ ,  $|E(G_p)| \geq \frac{ndp}{2} - \frac{\varepsilon n}{4} = \frac{Cn}{2} - \frac{\varepsilon}{4}n$ .

Let  $H$  be the subgraph of  $G_p$  induced by  $V \setminus V_0$ . Note that for every  $v \in V(G)$ , we have that  $d_H(v) < (1 + \delta)C$ . Hence, by Vizing's theorem [19], there exists a proper colouring of  $H$  with  $(1 + \delta)C$  colours. Therefore, with probability  $1 - o_n(1)$  there is a matching in  $H$  (and thus in  $G_p$ ) covering at least

$$\frac{2|E(H)|}{(1 + \delta)C} = \frac{2(|E(G_p)| - |E_0|)}{(1 + \delta)C} \geq \frac{(C - \varepsilon)}{(1 + \delta)C}n \geq (1 - \varepsilon)n$$

vertices, as required.  $\square$

## 2.2 Finding a nearly regular, nearly spanning subgraph

In this section, we prove Theorem 2. For ease of presentation, we write  $G := Q^d$ , so that  $G_p = Q_p^d$ . Throughout the section, we assume  $d \in \mathbb{N}$  is sufficiently large, and all asymptotic notation in this section will be with respect to the parameter  $d$ . Further, throughout this section, we let  $\delta := \delta(d)$  be a function tending to 0 arbitrarily slowly as  $d$  tends to infinity. We recall our assumption that  $dp \geq \log^5 d$ , and in particular,  $dp = \omega(1)$  (we will use this fact at various points in the proof).

We will analyse the following ‘pruning’ process on  $G_p$ . For each  $t \in \mathbb{N}$  let

$$\delta_t := \frac{t \cdot \delta}{\lfloor \log d \rfloor}.$$

Let  $\tau := \lfloor \log d \rfloor$  so that  $\delta_\tau = \delta$ . Let  $H_1 := G_p$ , and let

$$A_1 := \{v \in V(H_1) : d_{H_1}(v) \neq (1 \pm \delta_1)dp\}. \quad (1)$$

We then proceed as follows: At the  $t$ -th iteration, for  $t \in [2, \tau] \cap \mathbb{N}$ , we let  $H_t := G_p \left[ V(G) \setminus \bigcup_{i=1}^{t-1} A_i \right]$ , and let

$$A_t := \{v \in V(H_t) : d_{H_t}(v) < (1 - \delta_t)dp\}.$$

We run this process until  $t = \tau$ , and let

$$H := H_\tau \quad \text{and} \quad A := \bigcup_{t=1}^{\tau} A_t.$$

Note that if for some  $1 \leq t \leq \tau$ , we have  $A_t = \emptyset$ , then the process stabilises, that is, for every  $t' \geq t$ , we would have that  $A_{t'} = \emptyset$  and  $H_{t'} = H_t$ . Let us note here that the idea of iterative pruning for creating and analysing combinatorial structures is quite common in probabilistic combinatorics, see for example [3, 7, 18].

We will make use of the following observations:

- (i) The sets  $\{A_1, \dots, A_\tau\}$  are pairwise disjoint;
- (ii) If  $A_\tau = \emptyset$ , then for every  $v \in V(G) \setminus A$ ,  $d_H(v) = (1 \pm \delta)dp$ .

The first one is apparent by construction. To see (ii) assume that  $A_\tau = \emptyset$ . Then, every  $v \in (V(G) \setminus A) = \left(V(G) \setminus \bigcup_{t=1}^{\tau-1} A_t\right) = V(H_\tau)$  satisfies  $d_H(v) \geq (1 - \delta_\tau)dp = (1 - \delta)dp$ . On the other hand, since  $v \in (V(G) \setminus A) \subseteq (V(G) \setminus A_1)$ , it follows that  $d_H(v) \leq d_{G_p}(v) \leq (1 + \delta_1)dp \leq (1 + \delta)dp$ .

In particular, by (ii), in order to prove Theorem 2, it suffices to show that **whp**<sup>2</sup>

$$A_\tau = \emptyset \quad \text{and} \quad |A| = o(2^d).$$

Key to the proof is the following lemma, which gives a simple-to-analyse condition for when  $v \in A_t$  for  $t \in [2, \tau] \cap \mathbb{N}$ .

**Lemma 5.** *Let  $t \in [2, \tau] \cap \mathbb{N}$ . If  $v \in A_t$ , then there is a set  $X$  of at least  $\left(\frac{\delta dp}{2t \log d}\right)^{t-1}$  vertices at distance exactly  $t - 1$  from  $v$ , such for all  $x \in X$ ,*

$$d_{G_p}(x) \neq (1 \pm \delta_1)dp.$$

*Proof.* We will use the following observation about the structure of  $Q^d$ , which is easy to verify:

If  $k \leq d$  and  $v$  and  $w$  are at distance  $k$  in  $Q^d$ , then  $w$  has precisely  $d - k$  neighbours at distance  $k + 1$  from  $v$  and  $k$  neighbours at distance  $k - 1$  from  $v$ . (2)

The lemma will follow from iteratively applying the next fairly simple claim.

**Claim 6.** *Let  $t \in [2, \tau] \cap \mathbb{N}$  and  $k \in [1, t] \cap \mathbb{N}$ , let  $v \in V(G)$ , and let  $S \subseteq (A_t \cap N^k(v))$ . Then, there exists a set  $X \subseteq (A_{t-1} \cap N^{k+1}(v))$  with  $|X| \geq |S| \frac{\delta dp}{2(k+1) \log d}$ .*

*Proof.* Since each  $s \in S$  is in  $A_t$ ,  $d_{H_t}(s) < (1 - \delta_t)dp$ . On the other hand, since  $A_t$  and  $A_{t-1}$  are disjoint, each  $s \in S$  is not in  $A_{t-1}$  and consequently  $d_{H_{t-1}}(s) \geq (1 - \delta_{t-1})dp$ . Thus, there are at least  $(\delta_t - \delta_{t-1})dp = \frac{\delta dp}{\lfloor \log d \rfloor} \geq \frac{\delta dp}{\log d}$  neighbours of  $s$  which are in  $A_{t-1}$ , let us denote them by  $Y_s$ .

If we let  $X_s := (Y_s \cap N^{k+1}(v)) \subseteq (A_{t-1} \cap N^{k+1}(v))$ , then by (2),  $|X_s| \geq \frac{\delta dp}{\log d} - k$ . On the other hand, since  $X := \bigcup_{s \in S} X_s \subseteq N^{k+1}(v)$  and  $S \subseteq N^k(v)$ , again by (2) each  $x \in X$  lies in at most  $k + 1$  sets  $X_s$ . It follows by a simple double counting argument that

$$|X| \geq \frac{|S|}{k+1} \left( \frac{\delta dp}{\log d} - k \right) \geq |S| \frac{\delta dp}{2(k+1) \log d},$$

where we used that  $k \leq \tau = \lfloor \log d \rfloor \leq \log d$  and  $dp \geq \log^5 d$ . □

To complete the proof of Lemma 5, note that if  $v \in A_t$ , then  $v \notin A_{t-1}$ , and by the same argument as above, there exists  $S_1 \subseteq (N(v) \cap A_{t-1})$  with  $|S_1| \geq \frac{\delta dp}{2 \log d}$ .

Iteratively applying Claim 6 to the sets  $S_i$ , for  $i \in [1, t-2] \cap \mathbb{N}$ , we obtain a sequence of sets  $S_{i+1} \subseteq (N^{i+1}(v) \cap A_{t-i+1})$  with  $|S_{i+1}| \geq |S_i| \frac{\delta dp}{2(i+1) \log d}$ . It follows that

$$X := S_{t-1} \subseteq (N^{t-1}(v) \cap A_1)$$

has size at least

$$\prod_{i=1}^{t-1} \frac{\delta dp}{2(i+1) \log d} = \frac{1}{t!} \left( \frac{\delta dp}{2 \log d} \right)^{t-1} \geq \left( \frac{\delta dp}{2t \log d} \right)^{t-1},$$

where we used  $t! \leq t^{t-1}$ . Since  $X \subseteq A_1$ , by (1)  $X$  satisfies the assertion of the lemma. □

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

With these lemmas at hand, we are now ready to prove Theorem 2.

*Proof of Theorem 2.* It suffices to show that **whp**  $A_\tau = \emptyset$  and  $|A| = o(2^d)$ .

Fix  $v \in V(G)$  and  $t \in [2, \tau] \cap \mathbb{N}$ . We start by showing

$$\mathbb{P}[v \in A_t] \leq \exp \left\{ - \left( \frac{\delta dp}{2t \log d} \right)^{t-1} \right\}. \quad (3)$$

Indeed, by Lemma 5, if  $v \in A_t$ , then there is a set  $X$  of at least  $\left( \frac{\delta dp}{2t \log d} \right)^{t-1}$  vertices at distance  $t-1$  from  $v$ , such that  $d_{G_p}(x) \neq (1 \pm \delta_1)dp$  for all  $x \in X$ . Furthermore, since every  $x \in X$  is at distance exactly  $t-1$  from  $v$ , they have the same parity, and thus  $X$  is an independent set in  $G$ . For each  $x \in X$ ,  $d_{G_p}(x) \sim \text{Bin}(d, p)$  and so by Lemma 3, we have that

$$\mathbb{P}[d_{G_p}(x) \neq (1 \pm \delta_1)dp] \leq 2 \exp \left\{ - \frac{\delta_1^2 d^2 p^2}{3dp} \right\} \leq \exp \left\{ - \frac{\delta^2 dp}{4 \log^2 d} \right\}.$$

Since  $X \subseteq N^{t-1}(v)$ , there are at most  $\binom{|N^{t-1}(v)|}{|X|}$  possible choices for  $X$ . Note that  $|N^{t-1}(v)| = \binom{d}{t-1}$  (choosing  $t-1$  of the  $d$  coordinates to obtain a vertex at distance  $t-1$  from  $v$ ). Recalling that  $d_{G_p}(x)$  is independent for each  $x \in X$  and that  $|X| = r \geq \left( \frac{\delta dp}{2t \log d} \right)^{t-1}$ , by a union bound we obtain

$$\begin{aligned} \mathbb{P}[v \in A_t] &\leq \binom{\binom{d}{t-1}}{r} \exp \left\{ -r \cdot \frac{\delta^2 dp}{4 \log^2 d} \right\} \leq \left( (ed)^{t-1} \exp \{ -3 \log^2 d \} \right)^r \\ &\leq \exp \{ r (2 \log^2 d - 3 \log^2 d) \} \leq \exp \{ -2r \} \\ &\leq \exp \left\{ - \left( \frac{\delta dp}{2t \log d} \right)^{t-1} \right\}, \end{aligned}$$

where the third inequality follows since  $t \leq \tau = \lfloor \log d \rfloor \leq \log d$ . Note that in this estimation, we used (very generously) our assumption of  $p \geq \frac{\log^5 d}{d}$ , that is, that the numerator is polylogarithmic in  $d$ .

We first note that (3) implies that for each  $t \in [2, \tau] \cap \mathbb{N}$ ,

$$\mathbb{E}[|A_t|] \leq 2^d \exp \left\{ - \left( \frac{\delta dp}{2t \log d} \right)^{t-1} \right\} \leq 2^d \exp \{ -\log^2 d \}.$$

Also, by a similar application of Lemma 3, we have

$$\mathbb{E}[|A_1|] \leq 2^d \exp \left\{ - \frac{\delta^2 dp}{4 \log^2 d} \right\} \leq 2^d \exp \{ -\log^2 d \}.$$

Recalling that  $A := \bigcup_{t=1}^\tau A_t$ , it follows that  $\mathbb{E}[|A|] \leq 2^d \tau \exp \{ -\log^2 d \} = o(2^d)$ , and so by Markov's inequality, **whp**  $|A| = o(2^d)$ .

Secondly, by (3) we obtain

$$\begin{aligned} \mathbb{P}[A_\tau \neq \emptyset] &\leq \sum_{v \in V(G)} \mathbb{P}[v \in A_\tau] \leq 2^d \exp \left\{ - \left( \frac{\delta dp}{5 \log^2 d} \right)^{\tau-1} \right\} \\ &\leq 2^d \exp \left\{ - (\log^2 d)^{\lfloor \log d \rfloor - 1} \right\} \leq 2^d \exp \{ -d \} = o(1). \end{aligned}$$

□

### 3 Discussion

We showed that for any  $d$ -regular graph  $G$  on  $n$  vertices, for every  $\epsilon > 0$  there exists a constant  $C > 0$  such that if  $d \geq C$  and  $p \geq \frac{C}{d}$ , then there typically is a matching on at least  $(1 - \epsilon)n$  vertices in  $G_p$ . Further, we showed that when  $p \geq \frac{\log^5}{d}$ ,  $Q_p^d$  typically contains an induced nearly spanning subgraph, whose degrees are tightly concentrated around the expected degree  $dp$ . To find this subgraph, we employed a fairly simple pruning process.

In recent years there has been an interest in the universality of properties of  $G(n, p)$ , and in particular in extending results on the quantitative similarity of the structure of  $G(n, p)$  and  $Q_p^d$  to broader classes of *high-dimensional* graphs. For example, the typical emergence of the giant component, its uniqueness and its asymptotic properties have been considered [6, 8, 9, 10, 16]. Moreover, Diskin and Geisler [11] extended the result of Bollobás [4] to these settings as well, roughly showing that for any  $d$ -regular high-dimensional product graph, the hitting times of minimum degree one, connectivity, and perfect matching are **whp** the same. It is not hard to verify that the proof laid out in this paper generalises, almost verbatim, to  $d$ -regular high-dimensional Cartesian product graphs (specifically, the  $t$ -dimensional product of regular graphs of bounded order). In fact, one can even further relax the assumptions.

**Theorem 7.** *Let  $G$  be a  $d$ -regular graph with  $d = \omega(1)$ . Suppose that for every  $v \in V(G)$ , every  $k \leq \log d$ , and every  $u \in N^k(v)$ , we have that  $|N(u) \cap \bigcup_{i=1}^k N^i(v)| = O(\log d)$ . Let  $p \geq \frac{\log^5 d}{d}$ . Then, **whp**  $G_p$  contains an induced subgraph  $H$  such that  $|V(H)| = (1 - o_d(1))|V(G)|$  and for every  $v \in V(H)$ ,  $d_H(v) = (1 \pm o_d(1))dp$ .*

This raises the following more general question.

**Question 8.** *Let  $G$  be a  $d$ -regular graph with  $d = \omega(1)$ . What ‘minimal’ assumptions on  $G$  and  $p$  suffice to have: **whp**  $G_p$  has an induced nearly spanning, nearly regular subgraph with all degrees concentrated around  $dp$ ?*

As an application of Theorem 1 we have that when  $p \geq \frac{C}{d}$ , **whp** a largest matching in  $Q_p^d$  contains  $(1 - o_C(1))2^d$  vertices, that is, the first order term is  $2^d$  whereas the second order term is bounded by  $o_C(2^d)$ . Much more precise results are known in the setting of  $G(n, p)$  [13]. It would be interesting to determine more precisely the typical size of a largest matching in  $Q_p^d$  when  $p = \frac{C}{d}$  (for  $d > C$ ). A first step to answer this would be to determine the second order term — there, it is perhaps natural to conjecture that the size of the ‘defect set’ (that is, the set of vertices that are not covered by a largest matching in  $Q_p^d$ ) is dominated by the number of isolated vertices and thus **whp** a largest matching typically covers  $2^d - (2(1 - p))^d + o_C((2(1 - p))^d)$  vertices (indeed, this would resonate with the known hitting time results [4]).

**Acknowledgement** The second and third authors were supported in part by the Austrian Science Fund (FWF) [10.55776/P36131, 10.55776/F1002]. Part of this work was done while the first and the fourth authors were visiting Graz University of Technology, and they would like to thank the Graz University of Technology for its hospitality. For the purpose of open access, the authors have applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

### References

- [1] N. Alon, M. Krivelevich, and B. Sudakov. Large nearly regular induced subgraphs. *SIAM Journal on Discrete Mathematics*, 22(4):1325–1337, 2008.
- [2] N. Alon and J. H. Spencer. *The probabilistic method*. Hoboken, NJ: John Wiley & Sons, fourth edition, 2016.

- [3] J. Balogh, B. Bollobás, and R. Morris. Graph bootstrap percolation. *Random Structures Algorithms*, 41(4):413–440, 2012.
- [4] B. Bollobás. Complete matchings in random subgraphs of the cube. *Random Structures Algorithms*, 1(1):95–104, 1990.
- [5] B. Bollobás and A. Thomason. Random graphs of small order. In *Random graphs '83 (Poznań, 1983)*, volume 118 of *North-Holland Math. Stud.*, pages 47–97. North-Holland, Amsterdam, 1985.
- [6] M. Collares, J. Doolittle, and J. Erde. The evolution of the permutahedron. *arXiv:2404.17260*, 2024.
- [7] M. Collares, J. Erde, A. Geisler, and M. Kang. Universal behaviour of majority bootstrap percolation on high-dimensional geometric graphs. *arXiv:2406.17486*, 2024.
- [8] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Percolation on high-dimensional product graphs. *arXiv:2209.03722*, 2022.
- [9] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Percolation on irregular high-dimensional product graphs. *Combin. Probab. Comput.*, 33(3):377–403, 2024.
- [10] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Isoperimetric inequalities and supercritical percolation on high-dimensional graphs. *Combinatorica*, to appear, doi:10.1007/s00493-024-00089-0.
- [11] S. Diskin and A. Geisler. Perfect matching in product graphs and in their random subgraphs. *arXiv:2404.14020*, 2024.
- [12] P. Erdős and A. Rényi. On the existence of a factor of degree one of a connected random graph. *Acta Math. Acad. Sci. Hungar.*, 17:359–368, 1966.
- [13] A. Frieze. On large matchings and cycles in sparse random graphs. *Discrete Mathematics*, 59(3):243–256, 1986.
- [14] R. M. Karp and M. Sipser. Maximum matching in sparse random graphs. In *Proceedings of the 22nd Annual Symposium on Foundations of Computer Science*, pages 364–375, 1981.
- [15] M. Krivelevich, B. Sudakov, and N. Wormald. Regular induced subgraphs of a random graph. *Random Structures Algorithms*, 38(3):235–250, 2011.
- [16] L. Lichev. The giant component after percolation of product graphs. *J. Graph Theory*, 99(4):651–670, 2022.
- [17] D. Mitsche, M. Molloy, and P. Prałat.  $k$ -regular subgraphs near the  $k$ -core threshold of a random graph. *Journal of Combinatorial Theory, Series B*, 142:106–143, 2020.
- [18] B. Pittel, J. Spencer, and N. Wormald. Sudden emergence of a giant  $k$ -core in a random graph. *J. Combin. Theory Ser. B*, 67(1):111–151, 1996.
- [19] V. G. Vizing. On an estimate of the chromatic class of a  $p$ -graph. *Diskret. Analiz*, 3:25–30, 1964.