

# Powers of Hamiltonian cycles in randomly augmented Dirac graphs—The complete collection

Sylwia Antoniuk<sup>1</sup> | Andrzej Dudek<sup>2</sup>  | Andrzej Ruciński<sup>1</sup>

<sup>1</sup>Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland

<sup>2</sup>Department of Mathematics, Western Michigan University, Kalamazoo, Michigan, USA

## Correspondence

Andrzej Dudek, Department of Mathematics, Western Michigan University, Kalamazoo, MI, USA.  
Email: [andrzej.dudek@wmich.edu](mailto:andrzej.dudek@wmich.edu)

## Funding information

Simons Foundation; Narodowe Centrum Nauki

## Abstract

We study the powers of Hamiltonian cycles in randomly augmented Dirac graphs, that is,  $n$ -vertex graphs  $G$  with minimum degree at least  $(1/2 + \varepsilon)n$  to which some random edges are added. For any Dirac graph and every integer  $m \geq 2$ , we accurately estimate the threshold probability  $p = p(n)$  for the event that the random augmentation  $G \cup G(n, p)$  contains the  $m$ -th power of a Hamiltonian cycle.

## KEYWORDS

Dirac graphs, powers of Hamiltonian cycles, random graphs

## 1 | INTRODUCTION

In this paper we continue the study of powers of Hamiltonian cycles in randomly augmented Dirac graphs initiated by Bohman, Frieze, and Martin in [2] and developed in [1, 3, 5, 8]. For  $m \in \mathbb{N}$  the  $m$ -th power  $F^m$  of a graph  $F$  is defined as the graph on the same vertex set whose edges join distinct vertices at a distance at most  $m$  in  $F$ . The  $m$ -th power of a path or a cycle will be often called an  $m$ -path or, respectively, an  $m$ -cycle. Note that the  $m$ -cycle on  $v$  vertices has  $vm$  edges, while the  $v$ -vertex  $m$ -path has  $vm - \binom{m+1}{2}$  edges.

A Hamiltonian cycle in a graph  $G$  is a cycle which passes through all vertices of  $G$ . Extending the celebrated result of Dirac [4], Komlós, Sárközy, and Szemerédi in [6, 7] proved that for large  $n$  every graph  $G$  with  $n$  vertices and minimum degree  $\delta(G) \geq \frac{k}{k+1}n$  contains the  $k$ -th power of a Hamiltonian cycle.

## 1.1 | Thresholds and overthresholds

For an integer  $k \geq 0$ , a sequence of  $n$ -vertex graphs  $G = G(n)$  is  $k$ -Dirac if  $\delta(G) \geq \left(\frac{k}{k+1} + \varepsilon\right)n$  for a constant  $\varepsilon > 0$  and all  $n \geq n_0$ , for some  $n_0 \in \mathbb{N}$ . For integers  $m > k \geq 1$ , a  $(k, m)$ -Dirac threshold was defined in [1, Def. 1.1] with  $\alpha = k/(k+1) + \varepsilon$  as, roughly, the smallest function  $d(n)$  such that with probability approaching 1, for every  $k$ -Dirac sequence of graphs  $G(n)$  and every  $p = p(n) \geq Cd(n)$ , where  $C = C(\varepsilon, k) > 0$ , the union  $G(n) \cup G(n, p)$  contains the  $m$ -th power of a Hamiltonian cycle. If existed, such a threshold function was denoted by  $d_{k,m}(n)$ . The main feature of that definition was that it was independent of the constant  $\varepsilon$  (or  $\alpha$ ) — only the multiplicative constants depended on  $\varepsilon$ .

For  $k = 0$ , the threshold  $d_0 = n^{-1}$  has been established in [2]. In [5] it was proved that  $d_{k,k+1}(n) = n^{-1}$  for all  $k \geq 1$ . This was substantially extended in [1] to cover many other pairs  $(k, m)$ . In particular, it turned out that  $d_{k,m}(n) = n^{-1}$  for all  $m \leq 2k + 1$ . (This result was independently obtained by Nenadov and Trujić in [8]). For  $m = 2$ , Böttcher, Parczyk, Sgueglia, and Skokan in [3] extended this result to a more general setting when  $\delta(G) \geq \alpha n$ , for all values of  $\alpha \in (0, 1)$ .

However, a vast majority of pairs  $(k, m)$  were still unaccounted for. In this paper we consider the case  $k = 1$  only and solve the problem for all values of  $m$ . Along the way it turned out, however, that we had to come up with a new notion of threshold.

Looking back at the results in [1] for just  $k = 1$ , it is perhaps surprising that only the values of  $m \in \{2, 3, 4, 5, 8\}$  are covered. Indeed, Theorem 1.2 in [1] established the bound  $d_{1,m}(n) \leq n^{-2/\ell}$  where  $\ell = \min\{\ell' \leq m : \ell' \geq (m - \ell')(m - \ell' + 1)\}$ . On the other hand, Theorem 1.3 therein yielded the lower bound  $d_{1,m}(n) \geq n^{-2/(\lfloor m/2 \rfloor + 1)}$ , and these bounds coincide only for  $m = 2, 3, 4$ . For  $m = 8$  they leave a gap  $n^{-2/5} \leq d_{1,8}(n) \leq n^{-1/3}$  which was closed by showing that  $d_{1,8}(n) = n^{-1/3}$  (see [1, Theorem 1.5]).

The case  $m = 5$ , treated only as a concluding remark in [1], was exceptional in that it escaped the notion of threshold as defined there. Indeed, in Claims 9.1 and 9.2 in [1] the exponent of  $n$  depends on  $\varepsilon$ . More precisely, for each  $\varepsilon > 0$  there exist two constants  $c' = c'(\varepsilon)$  and  $c'' = c''(\varepsilon)$  such that the transition from limiting probability zero to one takes place somewhere between  $n^{-1/2-c'}$  and  $n^{-1/2-c''}$ . Moreover, both  $c'$  and  $c''$  converge to 0 with  $\varepsilon \rightarrow 0$ . (For more discussion on thresholds for up to  $m = 10$ , see Section 6 II).

Back then viewed rather as an eccentricity, the case  $m = 5$  has now become an enlightening indication of the truth in full generality. Indeed, it turns out that for all  $m \notin \{2, 3, 4, 8\}$ , the threshold behavior is similar to that for  $m = 5$ . As our goal has been to determine a threshold function independent of  $\varepsilon$ , we are in need of a new notion of a threshold. So, one might view the four “typical” cases solved in [1, 5] as the tip of an iceberg which, however, in its depths looks quite different than what the tip might have suggested.

Inspired by the case  $m = 5$ , we now define a new type of threshold. For future references, we do it for all  $k \geq 1$ , although this paper is exclusively devoted to the classical Dirac case of  $k = 1$ . For integers  $m \geq 1$  and  $n \geq m + 2$ , the family (or property)  $\mathcal{C}_n^m$  consists of all  $n$ -vertex graphs  $G$  that contain the  $m$ -th power of a Hamiltonian cycle.

We say that a function  $d(n)$  is a  $(k, m)$ -Dirac overthreshold if

- (i) for every  $\varepsilon > 0$  there exists  $\mu > 0$  such that for all  $n$ -vertex graphs  $G$  with  $\delta(G) \geq \left(\frac{k}{k+1} + \varepsilon\right)n$  and all  $p = p(n) \geq d(n)n^{-\mu}$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( G \cup G(n, p) \in \mathcal{C}_n^m \right) = 1$$

and

(ii) for every real  $\mu > 0$ , there exists  $\varepsilon > 0$  and a sequence of  $n$ -vertex graphs  $G_\varepsilon = G_\varepsilon(n)$  with  $\delta(G_\varepsilon) \geq \left(\frac{k}{k+1} + \varepsilon\right)n$  such that for every  $p = p(n) \leq d(n)n^{-\mu}$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( G_\varepsilon \cup G(n, p) \in \mathcal{C}_n^m \right) = 0.$$

If exists, the  $(k, m)$ -Dirac overthreshold is denoted by  $\bar{d}_{k,m}(n)$ . Obviously, the existence of the overthreshold  $\bar{d}_{k,m}(n)$  excludes the existence of the  $(k, m)$ -Dirac threshold  $d_{k,n}(n)$ . The prefix “over” is meant to remind us that the abrupt change in behavior of the probability in question happens just “below” the function  $d(n)$ . The main difference between the  $(k, m)$ -Dirac threshold as defined in [1] and the new notion above is that now the dependence on  $\varepsilon$  is much more substantial. As a drawback, however, for any given  $\varepsilon$  we do not obtain a threshold in the classical sense, but a pair of functions bounding it from both sides.

## 1.2 | Main results

The thresholds obtained in [1] were related to the density of the so-called braid graphs (see Definition 4.2 and Figure 4). Given a graph  $G$ , let  $v_G = |V(G)|$  and  $e_G = |E(G)|$ . For a graph  $G$  with  $v_G > 1$ , by *density* and, respectively, *maximum density*, we mean

$$d_G = \frac{e_G}{v_G - 1} \quad \text{and} \quad m_G = \max_{H \subseteq G, v_H > 1} d_G.$$

The parameter  $d_G$  is often called 1-density.

Roughly speaking, for  $0 \leq r \leq \ell$ , a braid graph consists of an ordered collection of  $\ell$ -cliques such that any two consecutive cliques are joined by a structure called an  $r$ -bridge having exactly  $\binom{r+1}{2}$  edges (see Definition 4.1 and Figure 3). It was crucial for the construction used in the proof in [1] that

$$m = k\ell + r, \tag{1}$$

so that when  $k + 1$  braids are intertwined and all  $\ell$ -cliques at distance at most  $k$ , one from each braid, are fully connected to each other, each vertex has exactly  $k\ell + r = m$  neighbors ahead of itself (and, by symmetry, also behind itself). That is, such a structure forms an  $m$ -path.

The cases of  $k$  and  $m$  covered in [1] were, with the exception of  $(1, 8)$  and  $(2, 14)$ , exactly those for which one could find integers  $\ell \geq 2$  and  $r \geq 0$  with  $r(r + 1) \leq \ell$  and such that  $(k + 1)(\ell - 1) \leq m \leq k\ell + r$  (the R-H-S inequality is due to the monotonicity of  $m$ -path containment with respect to  $m$ ). The lower bound on  $\ell$  implied that the densest subgraph of a braid was an  $\ell$ -clique and, consequently, the exponent in the  $(k, m)$ -Dirac threshold equalled the negated reciprocal of the density  $d_{K_\ell}$  of the  $\ell$ -clique, that is,  $-2/\ell$ . The reason was that for

$p \geq n^{-2/\ell}$  there are in  $G(n, p)$  plenty of copies of the braid (cf. [1, Proposition 5.8], an immediate corollary of [1, Theorem 2.2]), a fact which was a crucial building block in the proof therein.

In this paper we allow the opposite case, namely,  $\ell < r(r+1)$  (for  $k=1$ ). It turns out that now the densest subgraph of a braid is the braid itself (see Proposition 4.3). Moreover, another corollary of [1, Theorem 2.2], namely, Proposition 4.4, implies that, again, there are plenty of copies of braids in  $G(n, p)$  with appropriately adjusted  $p$  (the exponent of  $n$  is now the negated reciprocal of the density of the braid).

We aim at finding an optimal value of  $\ell$  for which the (asymptotic) density of the braid determines the  $(1, m)$ -Dirac overthreshold. For this sake we introduce a pivotal quantity  $f_m(\ell)$ . For integers  $m \geq 2$  and  $\ell \in [m]$ , let

$$f_m(\ell) = \frac{1}{\ell} \left( \binom{\ell}{2} + \binom{m-\ell+1}{2} \right) = \ell + \frac{m^2+m}{2\ell} - m - 1.$$

Note that for  $k=1$ , (1) becomes  $m = \ell + r$ . Thus, for  $\ell \geq m/2$ ,  $f_m(\ell)$  equals the average number of edges adjacent to the vertices of the first  $\ell$ -clique in the braid and, consequently, it equals the asymptotic density of the braid as the number of cliques tends to infinity (see Section 4 for details).

Observe that for  $\ell > 0$

$$f_m(\ell) = \left( \sqrt{\ell} - \frac{\sqrt{2m^2+2m}}{2\sqrt{\ell}} \right)^2 + \sqrt{2m^2+2m} - m - 1 \geq \sqrt{2m^2+2m} - m - 1.$$

Thus, the function  $f = f_m$  has a unique global minimum on the real interval  $(0, m]$  at

$$\lambda_m = \frac{\sqrt{2m^2+2m}}{2}.$$

Hence, the convexity of  $f$  implies that for each integer  $\ell \in [m]$  we have  $f(\ell) \geq \min\{f(\lfloor \lambda_m \rfloor), f(\lceil \lambda_m \rceil)\}$ . Let

$$\ell_m \in \{\lfloor \lambda_m \rfloor, \lceil \lambda_m \rceil\} \quad \text{be such that} \quad f(\ell_m) = \min\{f(\lfloor \lambda_m \rfloor), f(\lceil \lambda_m \rceil)\}. \quad (2)$$

As a convention, when  $\lfloor \lambda_m \rfloor \neq \lceil \lambda_m \rceil$  but  $f(\lfloor \lambda_m \rfloor) = f(\lceil \lambda_m \rceil)$ , we always set  $\ell_m = \lfloor \lambda_m \rfloor$ . For example,  $\lambda_{20} = \sqrt{210}$ ,  $\lfloor \lambda_{20} \rfloor = 14$ ,  $\lceil \lambda_{20} \rceil = 15$ , and  $f(14) = 8 = f(15)$ , so we set  $\ell_{20} = 14$ . In general,  $\lambda_m \geq m/2$  for all  $m$ , and  $\ell_m \geq \lambda_m - 1 \geq m/2$  for  $m \geq 4$ .

We are ready to state the main results of this paper. From now on we abbreviate  $(1, m)$ -Dirac (over)threshold to just  $m$ -Dirac (over)threshold and use  $\bar{d}_m(n)$  for  $\bar{d}_{1,m}(n)$ .

**Theorem 1.1.** *For every integer  $m \geq 2$  and a real  $\mu > 0$ , there exists  $\varepsilon > 0$  and a sequence of  $n$ -vertex graphs  $G_\varepsilon = G_\varepsilon(n)$  with  $\delta(G_\varepsilon) \geq \left(\frac{1}{2} + \varepsilon\right)n$  such that for every  $p = p(n) \leq n^{-1/f(\ell_m)-\mu}$*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G_\varepsilon \cup G(n, p) \in \mathcal{C}_n^m) = 0.$$

Theorem 1.1 implies that, if an overthreshold exists,  $\bar{d}_m(n) \geq n^{-1/f(\ell_m)}$ . Below we provide an upper bound.

**Theorem 1.2.** *For all integers  $m$  and  $\ell$  such that  $m/2 \leq \ell \leq m - 1$  and  $\ell < (m - \ell)(m - \ell + 1)$  and for all  $\varepsilon > 0$  there exists  $\mu > 0$  such that for all  $n$ -vertex graphs  $G$  with  $\delta(G) \geq (1/2 + \varepsilon)n$  and all  $p = p(n) \geq n^{-1/f(\ell) - \mu}$*

$$\lim_{n \rightarrow \infty} \mathbb{P}(G \cup G(n, p) \in \mathcal{C}_n^m) = 1.$$

Theorem 1.2 implies that, if exists,  $\bar{d}_m(n) \leq n^{-1/f(\ell)}$ .

Remember that in [1, 5] the  $m$ -Dirac thresholds for  $m \in \{2, 3, 4, 8\}$ , as well as the 5-Dirac overthreshold, have been already established. The next result states that for all other  $m$ , except 6 and 9, the parameter  $\ell_m$  from Theorem 1.1 satisfies the restriction on  $\ell$  from Theorem 1.2.

**Proposition 1.3.** *For  $m = 7$  and any integer  $m \geq 10$  we have*

$$\ell_m < (m - \ell_m)(m - \ell_m + 1).$$

We defer the simple proof of Proposition 1.3 to Appendix A.

Theorems 1.1 and 1.2 together with Proposition 1.3 determine the  $m$ -Dirac overthreshold in almost all cases.

**Corollary 1.4.** *For  $m = 7$  and every integer  $m \geq 10$ ,*

$$\bar{d}_m(n) = n^{-1/f(\ell_m)}.$$

The remaining cases ( $m = 6$  and  $m = 9$ ) are treated separately in Theorem 1.5. It is worth mentioning that here the overthreshold differs from  $n^{-1/f(\ell_m)}$ .

**Theorem 1.5.** *We have*

$$\bar{d}_6(n) = n^{-4/9} \quad \text{and} \quad \bar{d}_9(n) = n^{-2/7}.$$

Thereby, we have obtained the complete collection of  $m$ -Dirac thresholds and overthresholds. They are summarized in Table 1.

## 2 | LOWER BOUND—PROOF OF THEOREM 1.1

The proof of Theorem 1.1 relies on two lemmas. The first one is a simple observation about random graphs.

**TABLE 1** The negated reciprocals  $\alpha_m$  of the exponents of  $n$  in the  $m$ -Dirac thresholds ( $m \in \{2, 3, 4, 8\}$ ) and overthresholds (all other  $m \geq 2$ ).

$m$	2 [5]	3 [1, 8]	4 [1]	5 [1]	6	7	8 [1]	9	$\geq 10$
$\alpha_m$	1	1	$\frac{3}{2}$	2	$\frac{9}{4}$	$\frac{13}{5}$	3	$\frac{7}{2}$	$f(\ell_m)$

**Lemma 2.1.** For each integer  $m \geq 2$ , let  $\ell_m$  be as defined in 2. For every  $\mu > 0$ , if

$$p \leq n^{-1/f(\ell_m)-\mu},$$

then a.a.s.  $G(n, p)$  contains no copy of  $K_{m+1}$ .

*Proof.* Clearly,  $\mathbb{P}(G(n, p) \supset K_{m+1}) = O(n^{m+1}p^{\binom{m+1}{2}}) = o(1)$ , whenever  $p = o(n^{-2/m})$ . Thus, all we need to show is the inequality  $f(\ell_m) \leq m/2$ . For  $m \in \{2, 3\}$  this can be checked by computing directly  $f(\lfloor \lambda_m \rfloor)$  and  $f(\lceil \lambda_m \rceil)$ . Indeed, we have  $\lambda_2 = \sqrt{3}$  and  $\lambda_3 = \sqrt{6}$ , so  $\ell_2 = \ell_3 = 2$  and  $f(\ell_2) = 1/2 < 1$ , while  $f(\ell_3) = 1 < 3/2$ . Assume now that  $m \geq 4$ . It can be easily verified (by solving a quadratic inequality) that  $f(\ell_m) \leq m/2$  is equivalent to  $\ell_m \geq m/2$ . Finally, recall that  $\ell_m \geq \frac{\sqrt{2m^2+2m}}{2} - 1 \geq m/2$  for  $m \geq 4$ .  $\square$

The proof of the second lemma, which is fully deterministic, will be given in Section 3. Recall that by an  $m$ -path we mean the  $m$ -th power of a path.

**Lemma 2.2.** Let  $m \geq 2$  and let  $P$  be an  $m$ -path with  $V(P) = A \cup B$ ,  $A \cap B = \emptyset$ , such that there are no  $m + 1$  consecutive vertices in  $P$  either all belonging to  $A$  or all belonging to  $B$ . Then,

$$|E(P[A])| + |E(P[B])| \geq f(\ell_m)|V(P)| - 2m^2,$$

where  $\ell_m$  is defined in 2.

Lemma 2.2 is the key novelty compared to [1] and might be of independent interest, as it provides structural information about a graph commonly arising in the study of randomly augmented graphs.

Let us now define this crucial graph. It has been first used in [2] to prove a lower bound on  $d_0$ , and then in many other papers in the same context, for example, [1, 3, 5, 8].

**Definition 2.3.** For  $\varepsilon > 0$  and even  $n$ , let  $G$  be an  $n$ -vertex complete bipartite graph with bipartition classes  $X \cup Y$ , where  $|X| = |Y| = n/2$ . Fix two subsets  $U \subset X$  and  $W \subset Y$  with  $|U| = |W| = \lfloor \varepsilon n \rfloor$ . The graph  $G_\varepsilon$  is obtained from  $G$  by adding two complete bipartite graphs with bipartitions  $(U, X \setminus U)$  and  $(W, Y \setminus W)$ .

Now we are ready to give a short proof of Theorem 1.1.

*Proof of Theorem 1.1.* Given  $m$  and  $\mu$ , let  $\varepsilon > 0$  be any constant such that  $\varepsilon \leq 1/6$  and

$$\frac{1}{\varepsilon} > \frac{6m^2}{(f(\ell_m))^2\mu} + \frac{6m^2}{f(\ell_m)}. \quad (3)$$

(Note that  $\varepsilon$  decreases when  $\mu$  does.)

Let  $G_\varepsilon$  be as in Definition 2.3. Let  $p \leq n^{-1/f(\ell_m)-\mu}$  and suppose that  $G_\varepsilon \cup G(n, p)$  contains the  $m$ -th power of a Hamiltonian cycle  $C$ . After removing the vertices of  $U \cup W$

from  $G_\epsilon$  (and thus from  $C$ ), we obtain a collection of at most  $2\lfloor \epsilon n \rfloor$   $m$ -paths in  $(X \setminus U) \cup (Y \setminus W)$ , hence at least one of them must be of length at least

$$\frac{n - 2\lfloor \epsilon n \rfloor}{2\lfloor \epsilon n \rfloor} \geq \frac{1}{2\epsilon} - 1 \geq \frac{1}{3\epsilon}.$$

Fix such a path  $P$  of length  $L = \lceil \frac{1}{3\epsilon} \rceil$  and consider its edgewise intersection with  $G(n, p)$ . Setting  $A = V(P) \cap X$  and  $B = V(P) \cap Y$ , by the construction of  $G_\epsilon$  and the removal of  $U \cup W$ ,

$$G(n, p) \cap P = P[A] \cup P[B].$$

Let  $M = \lceil f(\ell_m)L \rceil - 2m^2$ . By Lemma 2.2, either  $G(n, p) \supset K_{m+1}$  or there are no  $m + 1$  consecutive vertices in  $P$  either all belonging to  $A$  or all belonging to  $B$ , and thus

$$|E(P[A])| + |E(P[B])| \geq M.$$

The former event, by Lemma 2.1, does not hold a.a.s. The latter, in turn, implies the existence in  $G(n, p)$  of a subgraph with  $L$  vertices and  $M$  edges. The expected number of such subgraphs, abbreviating  $f = f(\ell_m)$  and using our bound on  $p$  and (3), can be crudely bounded from above by

$$\binom{\binom{L}{2}}{M} n^L p^M \leq O(n^{L-(1/f+\mu)M}) = O(n^{L-(1/f+\mu)(fL-2m^2)}) = O(n^{2m^2/f+2m^2\mu-\mu f/(3\epsilon)}) = o(1).$$

Hence, a.a.s.  $G_\epsilon \cup G(n, p)$  cannot contain such a path  $P$ , and thus a.a.s.

$$G_\epsilon \cup G(n, p) \notin C_n^m.$$

□

### 3 | PROOF OF LEMMA 2.2

This section is entirely devoted to the proof of our main lemma, namely, Lemma 2.2.

*Proof of Lemma 2.2.* The ordering of the vertices on  $P$  defines a partition of  $A \cup B$  into segments  $S_1, S_2, \dots, S_t$ , which consist of maximal sets of consecutive vertices of  $P$  lying entirely in  $A$  or entirely in  $B$ . Thus,  $A = S_1 \cup S_3 \cup \dots$  and  $B = S_2 \cup S_4 \cup \dots$ . We will first show that to bound the number of edges in  $P[A] \cup P[B]$  we can consider a different  $m$ -path  $P'$  on the same set of vertices, but with a new partition  $A' \cup B' = A \cup B$  and new segments  $S'_1, S'_2, \dots, S'_t$  which alternate between  $A'$  and  $B'$ , and such that

- (a) for all  $i = 1, \dots, t'$ ,  $|S'_i| \leq m$ , and
- (b) for all  $i = 1, \dots, t' - 1$ ,  $|S'_i| + |S'_{i+1}| \geq m$ .

Condition (a) implies that neither  $P'[A']$ , nor  $P'[B']$ , contains a copy of  $K_{m+1}$ , while condition (b) guarantees that all edges in  $E(P'[A']) \cup E(P'[B'])$  connect only vertices lying in consecutive segments of  $A'$  or consecutive segments of  $B'$ . The second condition simplifies the computation of the number of relevant edges as, given (b), the number of edges between segment  $S'_i$  and all segments lying to the right will depend only on the size of  $S'_{i+1}$ . In the end, we will bound the size of  $E(P[A]) \cup E(P[B])$  by the size of  $E(P'[A']) \cup E(P'[B'])$  minus a correcting additive term, compare (4). Then the latter, owing to properties (a) and (b), will be minimized by  $f(\ell_m)|V(P')| + O(1)$ , a quantity corresponding to the case when  $P'[A']$  and  $P'[B']$  have both a braid-like structure.

The path  $P'$  and the partition  $A' \cup B'$  are constructed from the path  $P$  and the partition  $A \cup B$  iteratively. In each step we can shift a number of vertices between the partition classes, thus changing the partition. Moreover, once we shift a vertex from one class to another, say we shift a vertex  $v$  from  $A'$  to  $B'$ , we delete all edges going from  $v$  to vertices in  $A'$ , and we join  $v$  with those vertices in  $B'$  which lie at distance at most  $m$  from  $v$  in  $P'$ . Another valid operation is shifting the first vertex  $u$  of a segment  $S'_{i+1}$  at the end of segment  $S_{i-1}$ . Here we do not change the partition, but we change the ordering of vertices in  $P'$  and we need to adjust the edges so that  $u$  is connected only with vertices in the same partition class and lying at a distance at most  $m$  from it in the new ordering.

*Initiation:* Initially  $P' = P$ ,  $A' = A$ ,  $B' = B$ , and  $S'_i = S_i$  for  $i = 1, 2, \dots, t$ . If  $|S'_1| > m$ , we leave the first  $m$  vertices in  $S'_1$  and shift the remaining ones to  $S'_2$ . Note that this operation changes the partition, as the shifted vertices move from  $A'$  to  $B'$ , but it cannot increase the number of edges. Indeed, if  $T$  is the segment we shift from  $S'_1$  to  $S'_2$ , then for  $k \leq \lceil |T|/2 \rceil$  the number of edges incident with the  $k$ -th vertex from the left in  $T$ , which we have to remove, is at least as large as the number of edges incident with the  $k$ -th vertex from the right, which have to add.

Now, let  $j$  be the largest index for which  $|S'_1| + |S'_2| + \dots + |S'_j| < m$ . If  $j = 0$  or  $j = 1$  we do nothing. If  $j \geq 2$  and, say,  $S'_j \subset A'$ , we merge all vertices from  $S'_1 \cup S'_2 \cup \dots \cup S'_j$  into one class and place it in  $A'$ . That is we redefine  $S'_1 := S'_1 \cup S'_2 \cup \dots \cup S'_j$  and  $S'_i := S'_{i+j-1}$ ,  $i \geq 2$ . Note that this will increase the number of edges in  $P'[A'] \cup P'[B']$  by at most  $(m-1)^2/4$ . Indeed, if  $x$  and  $y$  are the numbers of vertices of  $S'_1 \cup S'_2 \cup \dots \cup S'_j$  in, respectively,  $A'$  and  $B'$ , then, since  $x + y < m$ , after the change the number of edges increases by  $xy \leq x(m-1-x) \leq (m-1)^2/4$ . Note that after this initial step, regardless of what the value of  $j$  was, we have  $|S'_1| \leq m$ .

*Iteration:* We then scan  $P'$  from left to right and check whether the segments  $S'_i$ ,  $i = 1, \dots, t'$ , satisfy conditions (a) and (b). If not, then we move some of the vertices, changing the ordering of  $P'$  or changing the partition  $A' \cup B'$ . Suppose that the segments  $S'_1, S'_2, \dots, S'_{i-1}$  of  $P'$  satisfy (a) and (b). In view of the above, we may assume that  $i \geq 2$ . If  $S'_{i-1}$  was the last segment, then we are done and, if not, we proceed by considering four cases. In each case we make suitable modifications of  $P'$  and the partition  $V(P') = A' \cup B'$  and bound the change of the number of edges in  $P'[A'] \cup P'[B']$ . It turns out that it never increases except in case 3 where it can increase by at most  $(m-1)^2/4$ .

*Case 0:* If  $|S'_i| \leq m$  and  $|S'_{i-1}| + |S'_i| \geq m$  then segments  $S'_1, S'_2, \dots, S'_i$  satisfy (a) and (b) and we can move to the next segment  $S'_{i+1}$ .

*Case 1:* If  $|S'_i| > m$ , we leave the first  $m$  vertices unchanged and move the remaining ones into  $S'_{i+1}$ , thus changing the partition (as the moved vertices change the side, see



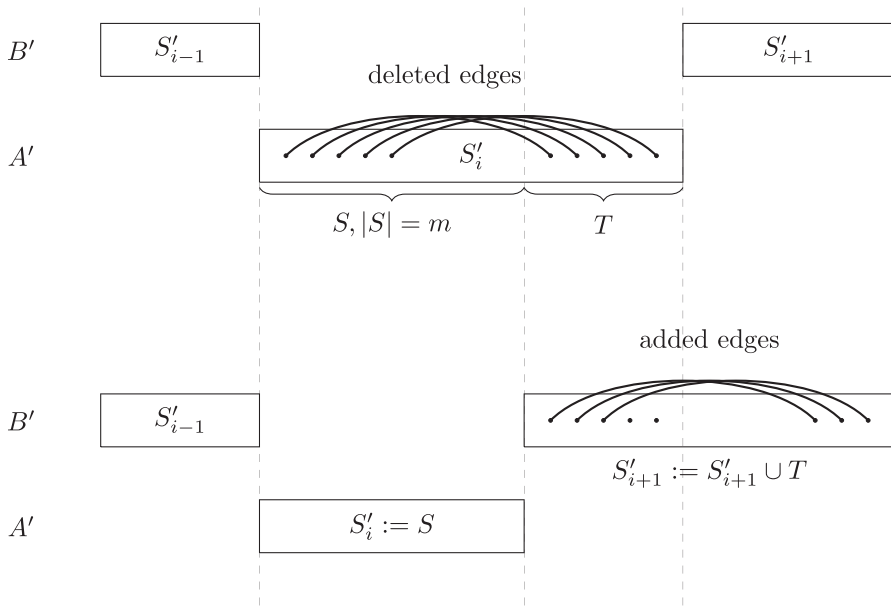


FIGURE 1 When  $|S'_i| > m$ , the last  $|S'_i| - m$  vertices of  $S'_i$  are shifted to  $S'_{i+1}$ .

Figure 1). If before this operation  $S'_i$  was the last segment, then we simply create a new segment  $S'_{i+1}$ . Notice that after this step segments  $S'_1, S'_2, \dots, S'_i$  satisfy (a) and (b), since now  $|S'_i| = m$ , and we can move to the next segment  $S'_{i+1}$ .

Let  $S$  be the sequence of the first  $m$  vertices of  $S'_i$  on  $P'$  before the change and let  $T$  be the sequence of the remaining vertices of  $S'_i$ . Notice that in this step we can only add or delete edges with exactly one endpoint in  $T$ . We delete edges between  $T$  and vertices on the left, and we add edges between  $T$  and some vertices on the right. Using the same argument as in the initiation we see that after this step the number of edges can either remain the same or drop.

Case 2: If  $|S'_{i-1}| + |S'_i| < m$  and  $S'_{i+1} \neq \emptyset$ , we shift the first vertex in  $S'_{i+1}$  to the end of  $S'_{i-1}$  (see Figure 2). Notice that after this operation segments  $S'_1, S'_2, \dots, S'_{i-1}$  still satisfy (a) and (b). Moreover, if before the shift  $|S'_{i+1}| = 1$  while  $S'_{i+2} \neq \emptyset$ , then, after the shift, we merge  $S'_i := S'_i \cup S'_{i+2}$  and renumber  $S'_h := S'_{h+2}$  for  $h \geq i + 1$ . We then check again if the assumption of Case 2 still holds for the new segments  $S'_{i-1}, S'_i, S'_{i+1}$ .

Without loss of generality, we may assume that  $S'_{i+1} \subset A'$ . Let  $u$  be the first vertex in  $S'_{i+1}$  (before the shift). Since the distance on  $P'$  between any two vertices  $v, w \in A' \setminus \{u\}$  does not change, the only edges of  $P[A']$  affected by the shift are those incident with  $u$ . Owing to the assumptions  $|S'_{i-2}| + |S'_{i-1}| \geq m$  when  $i > 2$ , and  $|S'_{i-1}| + |S'_i| < m$ , the number of neighbors of  $u$  to the left remains unchanged by the shift (in fact, it is precisely  $|S'_{i-1}|$ ). On the other hand, the distance on  $P'$  between  $u$  and any vertex in  $A'$  lying to the right increases (by  $|S'_i|$ ) after the shift, so the number of edges incident with  $u$  can only drop.

As for the edges in  $P[B']$  affected by the shift, we need to consider only those edges  $vw$  for which shifting  $u$  in front of  $S'_i$  has changed the distance between  $v$  and  $w$  on  $P'$ . Therefore, such edges need to be incident with  $S'_i$ . Notice that since  $|S'_{i-1}| + |S'_i| < m$ , before the shift each vertex in  $S'_i$  was adjacent to some vertex in  $S'_{i-2}$ , and, since

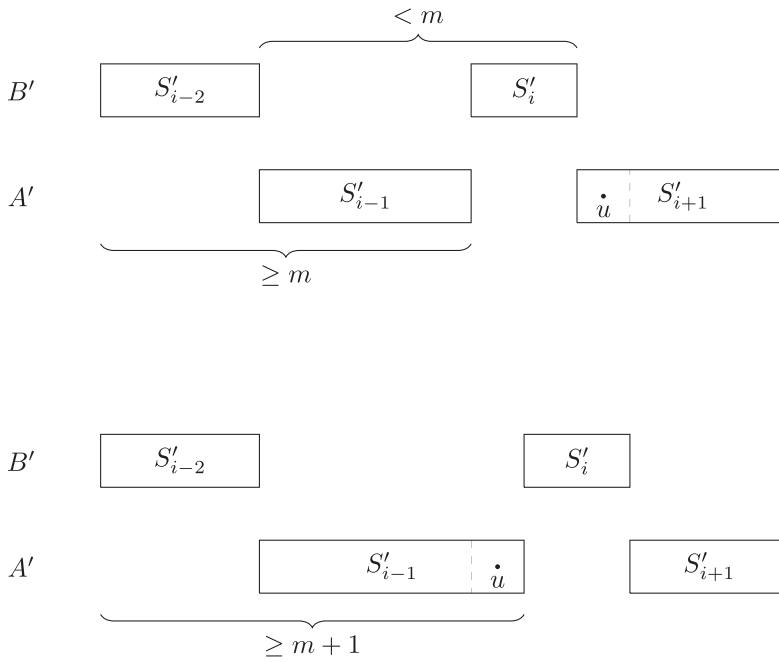


FIGURE 2 When  $|S'_{i-1}| + |S'_i| < m$ , the first vertex in  $S'_{i+1}$  is shifted at the end of  $S'_{i-1}$ .

$|S'_{i-2}| + |S'_{i-1}| \geq m$ , there were no other edges incident with  $S'_i$  and going to the left. Hence, after the shift the distance between  $S'_i$  and  $S'_{i-2}$  increases by 1 and the number of edges between  $S'_i$  and  $S'_{i-2}$  drops by  $|S'_i|$ . As for the edges incident with  $S'_i$  and going to the right, since after shifting  $u$  the distance between vertices in  $S'_i$  and vertices to the right decreases by one, we have to add at most  $|S'_i|$  new edges. Therefore, again, the total number of edges in  $P'[A'] \cup P'[B']$  can only drop. (The merging of  $S'_i$  and  $S'_{i+2}$  in the case when  $|S'_{i+1}| = 1$  does not alter any edges of  $P$ .)

*Case 3 (Termination):* If  $|S'_{i-1}| + |S'_i| < m$  and  $S'_{i+1} = \emptyset$ , we merge  $S'_{i-1} := S'_{i-1} \cup S'_i$  and finish the procedure. By doing so we have to add all edges between  $S'_{i-1}$  and  $S'_i$  and, similarly as in the initiation step, the number of such edges is at most  $(m - 1)^2/4$ .

Summing up, in the initiation step we added at most  $(m - 1)^2/4$  edges, then in each step we have not increased the number of edges until maybe at the termination where we added at most  $(m - 1)^2/4$  edges. Thus

$$|E(P[A])| + |E(P[B])| \geq |E(P'[A'])| + |E(P'[B'])| - \frac{(m - 1)^2}{2}. \tag{4}$$

It remains to bound the number of edges in  $P'[A'] \cup P'[B']$ . Suppose that the number of segments corresponding to the partition  $A' \cup B'$  is  $q$  and let  $x_1, x_2, \dots, x_q$  be the sizes of consecutive segments. We first count the number of edges in each segment. By (a) for each  $i = 1, 2, \dots, q$  we have  $x_i \leq m$ , hence each segment induces in  $P'[A'] \cup P'[B']$  exactly  $\binom{x_i}{2}$  edges. Next, we count the number of edges with endpoints in two different segments. By (b) for each  $i = 2, 3, \dots, q - 1$  we have  $x_{i-1} + x_i \geq m$  and  $x_i + x_{i+1} \geq m$ .

Therefore if we consider the edges incident with  $S'_{i-1}$  and going to the right, then each such edge can be incident only with  $S'_{i+1}$ . Moreover, the number of such edges is

$$(m - x_i) + (m - x_i - 1) + \dots + 1 = \frac{(m - x_i)(m - x_i + 1)}{2}.$$

Hence

$$\begin{aligned} |E(P'[A'])| + |E(P'[B'])| &= \sum_{i=1}^q \binom{x_i}{2} + \sum_{i=2}^{q-1} \frac{(m - x_i)(m - x_i + 1)}{2} \\ &\geq \sum_{i=1}^q \left( \binom{x_i}{2} + \frac{(m - x_i)(m - x_i + 1)}{2} \right) - m^2 \\ &= \sum_{i=1}^q x_i f(x_i) - m^2 \\ &\geq \sum_{i=1}^q x_i f(\ell_m) - m^2 = f(\ell_m)|V(P)| - m^2. \end{aligned}$$

Finally, by (4),

$$|E(P[A])| + |E(P[B])| \geq f(\ell_m)|V(P)| - 2m^2.$$

□

#### 4 | UPPER BOUND—PROOF OF THEOREM 1.2

The proof of Theorem 1.2 is similar to that of [1, Theorem 1.2] (with  $k = 1$ ) and therefore we will only discuss the necessary amendments. In turn, the proof of [1, Theorem 1.2] followed a general outline of the proof of [5, Theorem 1.1] which was based by nowadays standard method of absorption. We suggest the readers read Section 2 in [5] and Section 4 in [1] before delving into the rest of this section. In our proof all four pillars of the absorbing method: the Connecting Lemma, the Reservoir Lemma, the Absorbing Lemma, and the Covering Lemma, are exactly the same as, respectively, Lemmas 6.2, 4.2, 4.3, and 4.4 in [1], except for the assumption on  $p$ .

To describe this fundamental change with respect to the proof in [1], we begin by recalling the notions of a bridge and a braid graph which played a crucial role therein.

**Definition 4.1.** For  $r \geq 2$ , two sequences of vertices  $\vec{v} = (v_1, v_2, \dots, v_r)$  and  $\vec{u} = (u_1, u_2, \dots, u_r)$  of a graph  $G$  are said to form an  $r$ -bridge (or just a bridge if the value of  $r$  is clear from the context) if for each  $i = 1, 2, \dots, r$ , the vertex  $v_i$  is adjacent in  $G$  to all  $u_1, u_2, \dots, u_i$ .

**Definition 4.2.** For  $t \geq 1$ ,  $\ell \geq 2$ , and  $1 \leq r \leq \ell$ , let  $B(\ell, r, t)$  be the braid graph consisting of  $t$  vertex-disjoint  $\ell$ -cliques  $K_\ell^{(1)}, K_\ell^{(2)}, \dots, K_\ell^{(t)}$ , with vertices ordered arbitrarily, where for each  $i = 1, \dots, t - 1$ , the last  $r$  vertices of  $K_\ell^{(i)}$  and the first  $r$  vertices of  $K_\ell^{(i+1)}$  form an  $r$ -bridge. We write shortly  $B_t = B(\ell, r, t)$  and for any  $s \geq 1$ , we denote by  $sB_t$ , the union of  $s$  vertex disjoint copies of  $B_t$ .

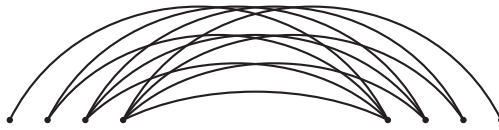


FIGURE 3 The 4-bridge.



FIGURE 4 The  $B(5, 3, 4)$  braid (consisting of four ordered cliques  $K_5$  joined by 3-bridges).

See Figures 3 and 4 for examples of a bridge and a braid.

As already mentioned above, the proof in [1] is based on the standard absorbing method whose four main ingredients, Connecting, Reservoir, Absorbing, and Covering Lemmas, all claim the existence of certain  $m$ -paths in  $G \cup G(n, p)$ , either of finite length  $2\ell t$  for some  $t \geq 2$ , or easily assembled from such. Each  $m$ -path of length  $2\ell t$  can be decomposed or embedded into the union of an  $\ell$ -blow-up  $P_{2t}(\ell)$  of a path of length  $2t$  and the pair of braids  $2B_t$ , where  $r = m - \ell$  (see [1, Proposition 5.5]). The proofs of the four crucial lemmas consist of a deterministic part and a probabilistic part, the former building several copies of  $P_{2t}(\ell)$  in  $G$ , the latter, based on [1, Proposition 5.8], a.a.s., complementing at least one of them with a suitable copy of  $2B_t$  (or its subgraph) coming from  $G(n, p)$ . And only the latter part of that proof has to be changed.

We now describe this change and its consequences. Recall the definitions of the density parameters  $d_G$  and  $m_G$  from Section 1.2. Clearly, as  $2B_t$  is a disjoint union of two copies of  $B_t$ , we have  $m_{2B_t} = m_{B_t}$ . In the course of the proof of [1, Proposition 5.8] it was shown that if  $\ell \geq r(r+1)$ ,  $r \geq 1$ , then for all  $t \geq 1$ , the maximum density in a braid is obtained by an  $\ell$ -clique, hence  $m_{B_t} = d_{K_\ell} = \ell/2$ . As a result, the assumption  $p \geq Cn^{-2/\ell}$  was sufficient in [1] to carry on the whole proof. Now, under the opposite assumption  $\ell < r(r+1)$ , the braid graph itself achieves the maximum density among all its subgraphs, that is,  $m_{B_t} = d_{B_t}$ .

**Proposition 4.3.** *For all  $t \geq 2$  and  $1 \leq r \leq \ell$  satisfying*

$$\ell < r(r+1),$$

*we have  $m_{B_t} = d_{B_t}$ .*

We defer the proof of Proposition 4.3 to the end of this section.

In view of Proposition 4.3, setting  $d_t = d_{B_t}$ , our assumption on  $p$  relaxes to  $p \geq Cn^{-1/d_t}$ . Most importantly,  $d_t$  is close to (but smaller than)  $f(\ell)$ . Indeed, recall that

$$f(\ell) = \frac{\binom{\ell}{2} + \binom{r+1}{2}}{\ell}.$$

On the other hand,

$$d_t = \frac{t \binom{\ell}{2} + (t-1) \binom{r+1}{2}}{t\ell - 1},$$

which can be rewritten as

$$d_t = f(\ell) - \frac{(\ell - 1)(r(r + 1) - \ell)}{2\ell(t\ell - 1)}.$$

Hence, for  $\ell < r(r + 1)$ ,  $d_t$  is a strictly increasing function of  $t$  and

$$\lim_{t \rightarrow \infty} d_t = f(\ell). \tag{5}$$

Next, we prove an analog of [1, Proposition 5.8] which reflects the change in  $m_{B_t}$  stemming from Proposition 4.3. For a graph  $G$  with at least one edge, set

$$\Psi_G = n^{v_G} p^{e_G} \quad \text{and} \quad \Phi_G = \min_{H \subseteq G, e_H > 0} \Psi_H.$$

**Proposition 4.4.** *Let  $t \geq 1, r \geq 1, \ell < r(r + 1), C \geq 1$ , and  $p \geq Cn^{-1/d_t}$ . Further, let  $\tau > 0, s \geq 1, B = sB_t, F \subset B$ , and  $\mathcal{F}$  be a family of at least  $\tau n^{v_F}$  copies of  $F$  in  $K_n$ . Finally, let  $X$  be the number of copies of  $F$  belonging to  $\mathcal{F}$  which are present in  $G(n, p)$ . Then there exists a constant  $c_F > 0$  such that*

$$\mathbb{P}(X \leq \tau \Psi_F / 2) \leq \exp\{-\tau^2 c_F Cn\}.$$

*Proof.* In view of [1, Theorem 2.2], setting  $c_F = 4^{-e_F} / 8$ , it suffices to show that  $\Phi_F \geq Cn$ . First assume that  $F' \subset F$  is connected, in particular,  $F' \subset B$ . Since  $C \geq 1$ , we have  $np^{d_t} \geq C \geq 1$ . Thus,

$$\Psi_{F'} \geq \Phi_{B_t} = \min_{H \subset B_t, e_H > 0} n(np^{d_H})^{v_H-1} \geq n(np^{m_{B_t}}) = n(np^{d_t}) \geq Cn,$$

where the middle inequality follows, since  $v_H \geq 2$  and  $np^{d_H} \geq np^{d_t} \geq 1$ . On the other hand, if  $F' = F_1 \cup F_2 \subset F$ , where  $F_1, F_2$  are vertex disjoint and, say,  $F_1$  is connected, then, by the above bound applied to  $F_1$ ,

$$\Psi_{F'} = \Psi_{F_1} \Psi_{F_2} \geq Cn \Psi_{F_2} > \Psi_{F_2},$$

so a disconnected  $F'$  does not achieve the minimum in  $\Phi_F$ . In summary,  $\Phi_F \geq Cn$  and the conclusion follows by [1, Theorem 2.2]. □

We are now in a position to outline the proof of Theorem 1.2.

*Proof of Theorem 1.2 (outline).* We essentially repeat the entire proof of [1, Theorem 1.2] with  $k = 1$  in which all applications of [1, Proposition 5.8] are replaced with applications of Proposition 4.4.

The exponent in the threshold probability  $p$  is thus determined by the largest value of  $t$  with which we apply Proposition 4.4. In [1], Proposition 5.8 has been applied four times (to either  $2B_t$  itself or to one of its subgraphs), namely, in the proofs of Lemma 6.2 (Connecting Lemma) with  $t = 4$ , Lemma 4.2 (Reservoir Lemma) again with  $t = 4$ , Proposition 7.2 (part of the proof of the Absorbing Lemma 4.3) twice with  $t = 2$ , and, last but not least, Claim 8.2 (part of the proof of the Covering Lemma 4.4) with  $t = 2M$ , where  $M \geq 1/(4\ell\gamma^3)$  by inequality (13) in [1].

The parameter  $\gamma$ , which within the proof of the Covering Lemma alone, must satisfy only the restriction  $\gamma \leq \varepsilon/12$ , in the main proof of [1, Theorem 1.2], is subject to much stronger restrictions, stemming from the other three main lemmas. Therefore, it would be very cumbersome and, in fact, not necessary for us, to determine the dependence of  $M$  on  $\varepsilon$  explicitly. All we can say is that  $M \rightarrow \infty$  with  $\varepsilon \rightarrow 0$ .

To finish the proof, it thus suffices to assume that  $p \geq n^{-1/d_{2M} + \varepsilon}$  (note that such a  $p$  satisfies the assumption of Proposition 4.4), set  $\mu = 1/d_{2M} + \varepsilon - 1/f(\ell)$ , so that the above assumption on  $p$  becomes  $p \geq n^{-1/f(\ell) - \mu}$ , and observe that, by (5), we have  $\mu \rightarrow 0$  with  $\varepsilon \rightarrow 0$  (as then  $M \rightarrow \infty$ ).

Now, the proof of [1, Theorem 1.2] can be repeated mutatis mutandis.  $\square$

We conclude this section with the proof of Proposition 4.3.

*Proof of Proposition 4.3.* We are going to prove a slightly stronger statement, namely, that  $B_t$  is *strictly* balanced, that is, for all *proper* subgraphs  $H$  of  $B_t$  we have  $d_H < d_t$ . First, consider the case when  $\ell \in \{r, r + 1\}$ . Then,  $B_t = B(\ell, r, t)$  is  $P_v^r$ , the  $r$ th power of the  $v$ -vertex path,  $v = t\ell$ , and we have

$$d_{P_v^r} = \frac{vr - \binom{r+1}{r}}{v-1} = r - \frac{\binom{r}{2}}{v-1}.$$

Note that  $d_{P_v^r}$  is a strictly increasing function of  $v$ . Therefore, it is easy to see that powers of paths are strictly balanced. Indeed, a proper subgraph  $H$  of  $P_v^r$  is either the  $r$ th power of a path  $P_{v'}$ ,  $v' < v$ , or a *proper spanning* subgraph of the  $r$ th power of a path  $P_{v'}$ ,  $v' \leq v$ , so in each case  $d_H < d_{P_v^r}$ .

From now on assume that  $\ell \geq r + 2$ . Our proof is by induction on  $t$ . It is easy to check that  $B_1 = K_\ell$  is strictly balanced. Let  $t \geq 2$  and assume that  $B_{t-1}$  is strictly balanced. Recall that the vertex set  $V$  of  $B_t$  is split into  $t$  disjoint cliques  $K_\ell^{(1)}, \dots, K_\ell^{(t)}$  the vertices of which are ordered (say, from left to right). Set  $V_i = V(K_\ell^{(i)})$ ,  $i = 1, \dots, t$ .

Let  $H$  be a proper subgraph of  $B_t$  such that  $d_H = m_{B_t}$  and  $|V(H) \cap V_i|$  is minimal. First observe that  $|V(H) \cap V_i| > 0$ , as otherwise  $H \subset B_{t-1}$ , and by the induction assumption and the strict monotonicity of  $d_t$ , we have  $m_{B_t} = d_H \leq d_{t-1} < d_t$ . Let  $s$  be the largest index such that  $V(H) \not\supset V_s$ .

*Claim 4.5.* We have  $s \leq t - 1$ .

This claim will be proved in Appendix B using tedious but elementary calculations. We now show that we can transform  $H$  into a subgraph  $H'$  with  $d_{H'} = d_H$  and  $|V(H') \cap V_i| < |V(H) \cap V_i|$ , reaching a contradiction.

Clearly,  $H$  is an induced subgraph of  $B_t$ , so it is fully determined by its vertex set. Set  $X = V(H) \cap V_s$  and  $Y = V(H) \cap V_t$ . Without loss of generality, we assume that  $Y$  is the leftmost subset of  $V_t$ . Further, let  $W \subset Y$  be the rightmost subset of  $Y$  of size  $|W| = \min(|Y|, \ell - |X|) > 0$ , as  $|Y| = |V(H) \cap V_t| > 0$  and  $\ell - |X| > 0$  by the choice of  $s$ . Moreover, let  $W' \subset V_s \setminus X$  be the rightmost subset of  $V_s \setminus X$  with  $|W'| = |W|$ . We also set  $Z = Y \setminus W$ . We create  $H'$  from  $H$  by replacing  $W$  with  $W'$  (see Figure 5). Observe that since  $|V(H') \cap V_t| < |V(H) \cap V_t|$ , it remains to show that  $e(H') \geq e(H)$ .

We consider two cases with respect to the value of  $s$ . For a vertex  $v \in V$  and two disjoint subsets  $U$  and  $W$  of  $V$  we denote by  $e(U, W)$  the number of edges of  $B_t$  with one endpoint in  $U$  and the other in  $W$ , and we denote by  $\text{deg}(v, U)$  the number of edges of  $B_t$  of the form  $vu$ , where  $u \in U$ .

Case  $s = t - 1$ : Given the definition of  $W$  and  $W'$ , we have

$$e(H') - e(H) \geq e(W', X \cup Z) - e(W, X \cup Z).$$

We show that the latter is nonnegative. Let  $\vec{X}$  be the rightmost subset of  $V_s$  with  $|\vec{X}| = |X|$  and  $U$  be the rightmost subset of  $V_s \setminus \vec{X}$  of size  $|U| = |W'|$ . Then,

$$e(W', X \cup Z) \geq e(U, \vec{X} \cup Z),$$

because  $U$  is further (or at the same distance) from the bridge between  $V_{t-1}$  and  $V_t$  than  $W'$  was. By the same reasoning,

$$e(W, \vec{X} \cup Z) \geq e(W, X \cup Z).$$

It remains to show that

$$e(U, \vec{X} \cup Z) \geq e(W, \vec{X} \cup Z). \tag{6}$$

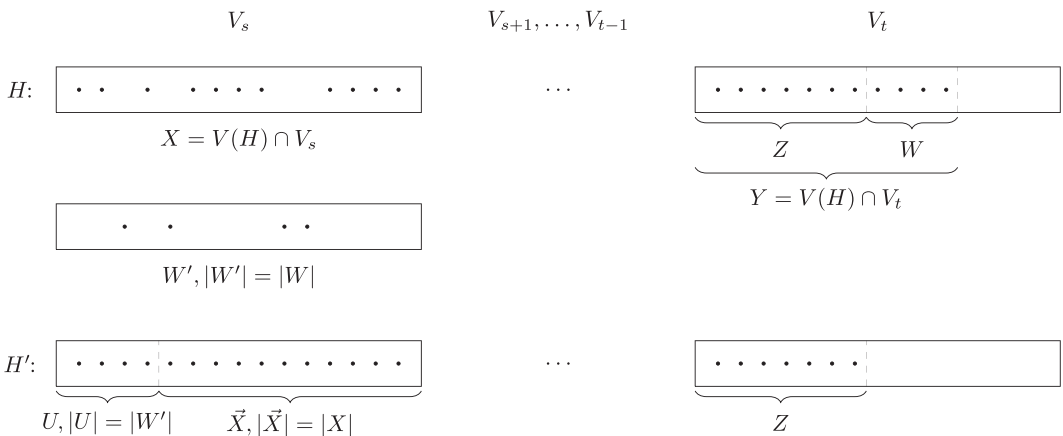


FIGURE 5 Creating  $H'$  from  $H$  by replacing  $W$  with  $W'$ . (Here  $|W| = \min(|Y|, \ell - |X|) = \ell - |X|$ )

To this end, set  $|U| = |W| = q$  and let  $U = \{u_{q-1}, \dots, u_0\}$  and  $W = \{w_0, \dots, w_{q-1}\}$ . Note that the indices of vertices (which follow the order imposed on  $V$ ) grow away from the bridge. Obviously,

$$e(U, \vec{X} \cup Z) = \sum_{j=0}^{q-1} \deg(u_j, \vec{X} \cup Z) \quad \text{and} \quad e(W, \vec{X} \cup Z) = \sum_{j=0}^{q-1} \deg(w_j, \vec{X} \cup Z).$$

Owing to the structure of the bridge, with  $x = |\vec{X}| = |X|$  and  $z = |Z|$ ,

$$\begin{aligned} \deg(u_j, \vec{X} \cup Z) &= \min(\max(x, r - j), x + z) \quad \text{and} \quad \deg(w_j, \vec{X} \cup Z) \\ &= \min(\max(z, r - j), x + z). \end{aligned}$$

Since  $x \geq z$ , we thus have  $\deg(u_j, \vec{X} \cup Z) \geq \deg(w_j, \vec{X} \cup Z)$ ,  $j = 0, \dots, q - 1$ , and (6) follows.

Case  $s \leq t - 2$ : Now

$$e(H') - e(H) \geq e(W', X \cup V_{s+1}) - e(W, Z \cup V_{t-1}),$$

and

$$e(W', X \cup V_{s+1}) \geq e(U, \vec{X} \cup V_{s+1}) \geq e(W, Z \cup V_{t-1}),$$

where the first inequality can be shown as above, and the second inequality follows from  $|X| \geq |Z|$ . This completes the proof.  $\square$

## 5 | PROOF OF THEOREM 1.5

To obtain the upper bounds on  $\bar{d}_6(n)$  and  $\bar{d}_9(n)$ , it is enough to apply Theorem 1.2 with, respectively,  $m = 6$ ,  $\ell = 4$ , and  $m = 9$ ,  $\ell = 6$ .

Let us then focus on the lower bounds. We are going to use a similar strategy as in the proof of Theorem 1.1, where the existence of the  $m$ -th power of a Hamiltonian cycle implied the existence of a too dense subgraph of  $G(n, p)$ . Here, however, we have to bound the number of edges more carefully.

The following terminology will be useful. Given an  $m$ -path  $P$  and a subset  $Z \subset V(P)$ , we say that  $u, v \in Z$  form a  $t$ -far  $ZZ$ -pair in  $P$  if there are in  $P$  exactly  $t - 1$  vertices from  $Z$  between  $u$  and  $v$ . If  $\{u, v\}$  happens to be an edge of  $P$ , we then call it a  $t$ -far  $ZZ$ -edge of  $P$ .

Case  $m = 6$ : Given  $\mu$ , let  $\varepsilon > 0$  be any constant such that  $\varepsilon \leq 1/12$  and

$$\frac{9}{16\varepsilon} > \frac{32}{9\mu} + 8. \quad (7)$$

Let  $G_\varepsilon$  be as in Definition 2.3,  $p = n^{-4/9-\mu}$ , and  $H = G_\varepsilon \cup G(n, p)$ . We define three events:

$\mathcal{C} - H$  contains the sixth power of a Hamiltonian cycle,



- $\mathcal{K}$  – the number of copies of  $K_5$  in  $G(n, p)$  is no more than  $n^{5/9}$ ,
- $\mathcal{S}$  – there is in  $G(n, p)$  a subgraph  $F$  with  $L = \left\lceil \frac{1}{4\epsilon} \right\rceil$  vertices and  $M = \left\lfloor \frac{9L}{4} \right\rfloor - 8$  edges.

Clearly,

$$\mathbb{P}(\mathcal{C}) \leq \mathbb{P}(\mathcal{C} \cap \mathcal{K}) + \mathbb{P}(\neg \mathcal{K}). \tag{8}$$

The main step of the proof will be to show that  $\mathcal{C} \cap \mathcal{K}$  implies  $\mathcal{S}$ . Taking this for granted, one can quickly complete the proof. Indeed, as  $p = o(n^{-4/9})$ , the expected number of copies of  $K_5$  in  $G(n, p)$  is less than  $n^5 p^{10} = o(n^{5/9})$ , and so, by Markov's inequality,  $\mathbb{P}(\neg \mathcal{K}) = o(1)$ . Similarly, using (7), the expected number of subgraphs of  $G(n, p)$  with  $L$  vertices and  $M$  edges can be bounded from above by

$$\binom{\binom{L}{2}}{M} n^L p^M = O\left(n^L n^{\binom{-4}{9} - \mu} \binom{9L}{4}^{-8}\right) = O\left(n^{\frac{32}{9} + 8\mu - \frac{9}{4}L\mu}\right) \leq O\left(n^{\frac{32}{9} + 8\mu - \frac{9}{16\epsilon}\mu}\right) = o(1).$$

Thus, another application of Markov's inequality yields  $\mathbb{P}(\mathcal{S}) = o(1)$ .

It remains to show the inclusion  $\mathcal{C} \cap \mathcal{K} \subset \mathcal{S}$ . Let  $C$  be the sixth power of a Hamilton cycle given by the event  $\mathcal{C}$ . After removing from  $C$  all vertices from  $U \cup W$ , as well as at least one vertex from each copy of at most  $n^{5/9} \leq \lfloor \epsilon n \rfloor$  copies of  $K_5$  in  $G(n, p)$ , we obtain a subgraph of  $C$  of order at least  $n - 3\lfloor \epsilon n \rfloor$ , which is a collection of at most  $3\lfloor \epsilon n \rfloor$  6-paths. Hence, at least one of them must have at least

$$\frac{n - 3\lfloor \epsilon n \rfloor}{3\lfloor \epsilon n \rfloor} \geq \frac{1}{3\epsilon} - 1 \geq \frac{1}{4\epsilon}$$

vertices. (Here, the last inequality follows from the assumption  $\epsilon \leq 1/12$ .)

Fix one such 6-path  $P$  on  $L = \left\lceil \frac{1}{4\epsilon} \right\rceil$  vertices which we relabel as  $V(P) = \{1, 2, \dots, L\}$  in the order of their appearance on  $P$ . Set  $A = V(P) \cap X$  and  $B = V(P) \cap Y$ , and notice that  $F = P[A] \cup P[B] \subset G(n, p)$  is  $K_5$ -free. We are going to derive a few observations regarding the structure of  $F$ .

First, observe that both  $P[A]$  and  $P[B]$  contain 2-paths as spanning subgraphs. Indeed, let  $u, v \in A$ ,  $u < v$ , be a  $t$ -far  $AA$ -pair,  $t \in \{1, 2\}$ . As  $P$  is a 6-path, it suffices to show that  $v - u \leq 6$ . Suppose for contradiction that  $v - u \geq 7$ . Then, there are at least five vertices in  $B$  between  $u$  and  $v$  on  $P$ , and the first five of them, say  $w_1, \dots, w_5$ , as they satisfy  $w_5 - w_1 \leq 6$ , induce a copy of  $K_5$  in  $P[B]$ , a contradiction. Hence,  $P[A]$  and (by symmetry)  $P[B]$  each contain a spanning 2-path.

Since the 2-path on  $s$  vertices has precisely  $2s - 3$  edges, we conclude that the number of 1-far and 2-far edges in  $F$  is

$$(2|A| - 3) + (2|B| - 3) = 2L - 6. \tag{9}$$

Next, we are going to bound the number of 3-far edges in  $F$ . Let  $R$  be a segment of any seven consecutive vertices from  $P$ . Since there is no  $K_5$  in  $F$ , either  $|R \cap A| = 4$  and  $|R \cap B| = 3$ , or  $|R \cap A| = 3$  and  $|R \cap B| = 4$ . So, there is a 3-far  $AA$ -edge or 3-far  $BB$ -edge within  $R$ . Since

there are  $L - 6$  consecutive segments  $R$  on seven vertices and a given 3-far edge may be contained in at most four such segments, we infer that the number of 3-far edges in  $F$  is at least

$$\frac{L - 6}{4} \geq \frac{L}{4} - 2. \quad (10)$$

Estimate (10), together with (9), implies that  $|F| \geq \left\lceil \frac{9L}{4} \right\rceil - 8 = M$ . Thus, the event  $\mathcal{S}$  holds which finishes the proof of Theorem 1.5 for  $m = 6$ .

Case  $m = 9$ : The proof is very similar to the previous case. Given  $\mu > 0$ , let  $p = n^{-2/7-\mu}$  and  $\varepsilon > 0$  satisfy

$$\frac{7}{8\varepsilon} > \frac{10}{7\mu} + 5. \quad (11)$$

Let again  $G_\varepsilon$  be as in Definition 2.3. This time we define events

- $\mathcal{C}$  -  $H$  contains the ninth power of a Hamiltonian cycle,
- $\mathcal{K}$  - the number of copies of  $K_7$  in  $G(n, p)$  is no more than  $n/\log n$ ,
- $\mathcal{S}$  - there is in  $G(n, p)$  a subgraph  $F$  with  $L = \left\lceil \frac{1}{4\varepsilon} \right\rceil$  vertices and  $M = \left\lceil \frac{7L}{2} \right\rceil - 5$  edges.

An analog of (8) still holds. The expected number of copies of  $K_7$  in  $G(n, p)$  is  $O(n^{1-27\mu}) = o(n/\log n)$ , so, by Markov's inequality,  $\mathbb{P}(\neg\mathcal{K}) = o(1)$ . Similarly, using (11), one can show that also  $\mathbb{P}(\mathcal{S}) = o(1)$ . It remains to show that  $\mathcal{C} \cap \mathcal{K}$  implies  $\mathcal{S}$ .

Suppose that  $G_\varepsilon \cup G(n, p)$  contains a subgraph  $C$  which is the ninth power of a Hamiltonian cycle. As before, after removing from  $G_\varepsilon \cup G(n, p)$  all vertices from  $U \cup W$  as well as at least one vertex from each copy of  $K_7$  in  $G(n, p)$ , we obtain a 9-path  $P$  on  $L = \left\lceil \frac{1}{4\varepsilon} \right\rceil$  vertices. Let  $A, B$ , and  $F$  be as before and let the vertices of  $P$  be relabeled as  $V(P) = \{1, \dots, L\}$ .

Benefiting from the absence of  $K_7$  in  $F$ , it can be shown, similarly to the case  $m = 6$ , that both  $P[A]$  and  $P[B]$  contain spanning 3-paths. Indeed, let  $u, v \in A$ ,  $u < v$ , be a  $t$ -far  $AA$ -pair,  $t \in \{1, 2, 3\}$ . Suppose that  $v - u \geq 7$ . Then, there are at least seven vertices in  $B$  between  $u$  and  $v$  on  $P$ , and, consequently, a clique  $K_7$  in  $P[B]$ , a contradiction. Hence,  $P[A]$  and (by symmetry)  $P[B]$  each contain a spanning 3-path. This yields  $3L - 12$   $t$ -far edges in  $F$  with  $t \leq 3$ . We are going to show that there are another  $L/2$  edges in  $F$  (4-far and 5-far).

Let  $R$  be a segment of any 10 consecutive vertices from  $P$ . Since there is no  $K_7$  in  $F$ ,  $R$  contains five vertices from each set  $A$  and  $B$ , or six vertices from one of them (and 4 from the other). In either case there are two 4-far edges within  $R$ . As there are  $L - 9$  segments of length 10 and each 4-far edge may belong to at most six of them, the total number  $w$  of 4-far edges in  $F$  satisfies

$$w \geq \frac{2(L - 9)}{6} = \frac{L}{3} - 3.$$

This is not quite enough, but there may also be 5-far edges in  $F$ . Let us denote their number by  $z$ . Note that a very small value of  $z$  can boost our bounds of  $w + z$  even higher than a large one. In

particular, it is easy to check that if  $z = 0$ , then  $w = L - 9$ . So, we have to optimize by relating  $w$  with  $z$ . Let  $w_A, w_B, z_A, z_B$  be the numbers of 4-far and 5-far edges in, respectively,  $P[A]$  and  $P[B]$ .

Let  $u < v$  be a 4-far  $AA$ -pair. The only reason for it not to be a 4-far edge is that  $v - u \geq 10$ , in which case there is a 5-far  $BB$ -edge  $u'v'$  with  $u < u' < v' < v$ . So, we can map every 4-far  $AA$ -pair which is not an edge of  $P[A]$  to a 5-far  $BB$ -edge. Note that in this mapping the pre-image of a fixed 5-far  $BB$ -edge has size at most 4. Thus,  $|A| - 4 - w_A \leq 4z_B$ , and, by a symmetric argument,  $|B| - 4 - w_B \leq 4z_A$ . Summing up, we get  $L - 8 - w \leq 4z$ , and, taking into account the previous bound on  $w$ , we have  $w \geq \max(L/3 - 3, L - 4z - 8)$ . In the end, we want to minimize the quantity

$$w + z \geq \max(L/3 - 3, L - 4z - 8) + z.$$

It is easy to check that the minimum is achieved at  $z = L/6 - 5/4$  and equals  $L/2 - 17/4 \geq L/2 - 5$ . Thus,  $F$  has in total at least  $3L + \lfloor \frac{L}{2} \rfloor - 5$  edges and the event  $S$  holds. This completes the proof of Theorem 1.5.

## 6 | CONCLUDING REMARKS

**I.** Recall that to determine the  $m$ -Dirac overthreshold  $\bar{d}_m$ , we established Theorems 1.1 and 1.2. For a fixed  $\varepsilon > 0$  and a given  $n$ -vertex graph  $G$  with  $\delta(G) \geq (1/2 + \varepsilon)n$ , these theorems yield bounds on the ordinary threshold probability  $\hat{p}$  for the property  $G \cup G(n, p) \in \mathcal{C}_n^m$  (assuming it exists), namely,

$$n^{-1/f(\ell)-\mu_1} \leq \hat{p} \leq n^{-1/f(\ell)-\mu_2},$$

where  $\mu_1 = \Theta(\varepsilon)$ , while  $\mu_2 = \mu_2(\varepsilon)$  is an implicit function of  $\varepsilon$  with  $\lim_{\varepsilon \rightarrow 0} \mu_2(\varepsilon) = 0$  — compare (3) and, respectively, the proof of Theorem 1.2. It would be nice, but probably extremely difficult, to close the gap.

**Problem 6.1.** For all integers  $2 \leq \ell \leq m - 1$  such that  $\ell < (m - \ell)(m - \ell + 1)$  and for all  $\varepsilon > 0$  find a function  $\mu = \mu(\varepsilon) > 0$  such that for all  $n$ -vertex graphs  $G$  with  $\delta(G) \geq (1/2 + \varepsilon)n$

$$\lim_{n \rightarrow \infty} \mathbb{P}(G \cup G(n, p) \in \mathcal{C}_n^m) = \begin{cases} 1 & \text{if } p \gg n^{-1/f(\ell)-\mu}, \\ 0 & \text{if } p \ll n^{-1/f(\ell)-\mu}. \end{cases}$$

(Here  $a_n \ll b_n$  means  $a_n = o(b_n)$ , while  $a_n \gg b_n$  means  $b_n = o(a_n)$ .)

**II.** One may wonder what is so special about the cases  $m = 2, 3, 4, 8$  that just for them we have obtained the standard Dirac-threshold  $d_m(n)$ , while for all other values of  $m$  we ended up with the Dirac overthreshold  $\bar{d}_m(n)$ . A humorous answer could be that it is yet another instance of the *Law of Small Numbers*: whatever seems to be true for small instances of a parameter, fails to hold in general. Still, in this subsection we attempt to provide a more intelligent answer.

TABLE 2 Summary of all thresholds for  $2 \leq m \leq 10$ .

$m$	$\ell$	$r$	$f(\ell)$	$f(\ell - 1)$	Thm 1.2 [1]	Thm 1.3 [1]	Thm 1.2	Thm 1.1	Dense structure in $G(n, p)$	Threshold
2	2	0	n/a	n/a	$\leq n^{-1}$	$\geq n^{-1}$	–	–	$\Theta(n)$ edges $\geq cn^{-1}$	$d_2 = n^{-1}$
3	2	1	n/a	n/a	$\leq n^{-1}$	$\geq n^{-1}$	–	–	$\Theta(n)$ edges $\geq cn^{-1}$	$d_3 = n^{-1}$
4	3	1	n/a	n/a	$\leq n^{-2/3}$	$\geq n^{-2/3}$	–	–	$\Theta(n)$ triangles $\geq cn^{-2/3}$	$d_4 = n^{-2/3}$
5	4	1	$\frac{7}{4}$	2	$\leq n^{-1/2}$	–	$\leq n^{-1/2-\mu}$	$\geq n^{-4/7-\mu}$	2-path $\geq n^{-1/2-\mu}$	$\bar{d}_5 = n^{-1/2}$
6	5	1	$\frac{11}{5}$	$\frac{9}{4}$	$\leq n^{-2/5}$	–	$\leq n^{-4/9-\mu}$	$\geq n^{-5/11-\mu}$	2-path, 3-far edges $\geq n^{-4/9-\mu}$	$\bar{d}_6 = n^{-4/9}$
7	6	1	$\frac{8}{3}$	$\frac{13}{5}$	$\leq n^{-1/3}$	–	$\leq n^{-5/13-\mu}$	$\geq n^{-5/13-\mu}$	not needed	$\bar{d}_7 = n^{-5/13}$
8	6	2	3	$\frac{16}{5}$	$\leq n^{-1/3}$	–	$\leq n^{-5/16-\mu}$	$\geq n^{-1/3-\mu}$	3-path, 4-far edges $\geq cn^{-1/3}$	$d_8 = n^{-1/3}$
9	7	2	$\frac{24}{7}$	$\frac{7}{2}$	$\leq n^{-2/7}$	–	$\leq n^{-2/7-\mu}$	$\geq n^{-7/24-\mu}$	3-path, 4,5-far edges $\geq n^{-2/7-\mu}$	$\bar{d}_9 = n^{-2/7}$
10	8	2	$\frac{31}{8}$	$\frac{27}{4}$	$\leq n^{-1/4}$	–	$\leq n^{-4/27-\mu}$	$\geq n^{-4/27-\mu}$	not needed	$\bar{d}_{10} = n^{-4/27}$

Note: Here  $\ell$  is the smallest integer satisfying  $\ell \geq r(r + 1)$ , where  $r = m - \ell$ ; circled fractions correspond to  $\ell_m$ ; shaded boxes indicate which theorems and ad hoc techniques (column “Dense structure in  $G(n, p)$ ”) determine the threshold.

Altogether, we have had at our disposal four statements and one ad hoc technique to deduce upper and lower bounds on the order of magnitude of  $p = p(n)$  which guarantees the presence of a Hamiltonian  $m$ -cycle in  $G \cup G(n, p)$ , where  $G$  is a Dirac graph: Theorems [1, Theorem 1.2] and 1.2 for upper bounds, Theorems [1, Theorem 1.3] and 1.1 for lower bounds, the latter sometimes strengthened by forcing some extra edges in  $G(n, p)$ .

For  $m = 2, 3, 4$ , the pair of theorems from [1] suffices to pinpoint the threshold  $d_m$ . In particular, the lower bounds follow by observing that a Hamiltonian  $m$ -cycle in  $G \cup G(n, p)$  must have a linear number of edges ( $m = 2, 3$ ) or triangles ( $m = 4$ ) in  $G(n, p)$  alone which is unlikely as  $p \leq cn^{-1}$  ( $m = 2, 3$ ) or  $p \leq cn^{-2/3}$  ( $m = 4$ ). For  $m = 8$ , [1, Theorem 1.3] does not apply, yet we have an ad hoc remedy: for  $p \leq cn^{-1/3}$  there are a.a.s. no copies of  $K_6$  in  $G(n, p)$  which implies that there have to be long 3-paths with a linear number of extra 4-edges which, again, is very unlikely.

With three exceptions, for all other  $m$  we have  $\ell_m < (m - \ell_m)(m - \ell_m + 1)$  (cf. Proposition 1.3) and the pair of Theorems 1.2 and 1.1 “squeezes out” an overthreshold  $\bar{d}_m$ . Why not  $d_m$ ? On the one hand, because for such  $\ell_m$  the corresponding braids are balanced (cf. Proposition 1.3) and, on the other hand, because  $p$  has to be small enough to “eliminate” dense paths from  $G(n, p)$  (cf. Lemma 2.2).

The three remaining cases ( $m = 5, 6, 9$ ), at least in principle, could go either way. Let us focus on  $m = 5$  as the other two are quite similar (though a bit more involved). We have  $\ell_5 = 4$  and  $f(\ell_5) = 7/4$ , while  $f(3) = 2$  is the runner-up. Thus, Theorems 1.2 and 1.1 imply, respectively, that  $\bar{d}_5(n) \leq n^{-1/2}$  and  $\bar{d}_5(n) \geq n^{-4/7}$ . At this point it is still possible that for some  $4/7 < \tau < 1/2$  we have a Dirac threshold  $d_5(n) = n^{-\tau}$ . However, it turned out that the

lower bound can be dramatically improved to match the upper bound and thus to establish  $\bar{d}_5(n) = n^{-1/2}$ . Again, this improvement was possible by analyzing the structure of  $G_\varepsilon \cup G(n, p)$  and showing that in  $G(n, p)$  alone either there are linearly many copies of  $K_4$  or a long 2-path, both quite unlikely when  $p \leq n^{1/2-\mu}$ . Finally, note that the overthresholds for  $m = 5, 6, 9$  are *not* of the form  $n^{-1/f(\ell_m)}$ .

For the convenience of the reader, we have summarized the above discussion in Table 2.

**III.** This paper is exclusively devoted to the classical Dirac case  $k = 1$ . However, as it was already mentioned in Introduction, in [1, Theorems 1.3 and 1.5] the authors determined the usual  $(k, m)$ -Dirac thresholds  $d_{k,m} = n^{-2/\ell}$  for every  $k \geq 1$  and all  $m$  falling into the interval  $(k+1)(\ell-1) \leq m \leq k\ell + (\sqrt{4\ell-1}-1)/2$  (for a fixed  $k$  there is only a finite number of feasible choices of  $m$ ), as well as  $d_{1,8} = d_{2,14} = n^{-1/3}$ . We believe that, as in the case  $k = 1$  (cf. discussion in Subsection II above), for all other pairs  $(k, m)$ , the  $(k, m)$ -Dirac thresholds  $d_{k,m}$  do not exist and should be replaced by the overthresholds  $\bar{d}_{k,m}$ .

**Problem 6.2.** Determine the overthreshold  $\bar{d}_{k,m}$  for all pairs  $(k, m)$ ,  $k \geq 2$ , not covered in [1].

We anticipate that the main difficulty in generalizing our results to arbitrary  $k \geq 1$  would be, again, to formulate and prove a  $k$ -analog of Lemma 2.2.

## ACKNOWLEDGMENTS

We are extremely grateful to both referees for their meticulous reading of the manuscript and numerous invaluable comments which have led to a substantial improvement of our submission. We would also like to thank Christian Reiher for quickly rebutting a conjecture we used to have with respect to Problem 6.2. The second author was supported in part by Simons Foundation Grant #522400. The third author was supported by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426.

## ORCID

Andrzej Dudek  <http://orcid.org/0000-0002-6410-5413>

## REFERENCES

1. S. Antoniuk, A. Dudek, C. Reiher, A. Ruciński, M. Schacht, *High powers of Hamiltonian cycles in randomly augmented graphs*, J. Graph Theory. **98** (2021), no. 2, 255–284.
2. T. Bohman, A. Frieze, and R. Martin, *How many random edges make a dense graph Hamiltonian?* Random Structures Algorithms. **22** (2003), no. 1, 33–42. MR1943857.
3. J. Böttcher, O. Parczyk, A. Sgueglia, and J. Skokan, *The square of a Hamilton cycle in randomly perturbed graphs*, arXiv:2202.05215 (2022).
4. G. A. Dirac, *Some theorems on abstract graphs*, Proc. Lond. Math. Soc. **2** (1952), no. 3, 69–81. MR47308.
5. A. Dudek, Chr. Reiher, A. Rucinski, M. Schacht, *Powers of Hamiltonian cycles in randomly augmented graphs*, Random Structures Algorithms. **56** (2020), no. 1, 122–141.
6. J. Komlós, G. N. Sárközy, and E. Szemerédi, *On the square of a Hamiltonian cycle in dense graphs*, Random Structures Algorithms. **9** (1996), no. 1–2, pp. 193–211.
7. J. Komlós, G. N. Sárközy, and E. Szemerédi, *Proof of the Seymour conjecture for large graphs*, Ann. Comb. **2** (1998), no. 1, 43–60. MR1682919.
8. R. Nenadov and M. Trujić, *Sprinkling a few random edges doubles the power*, SIAM J. Discrete Math. **35** (2021), no. 2, 988–1004. MR4259197.

**How to cite this article:** S. Antoniuk, A. Dudek, and A. Ruciński, *Powers of Hamiltonian cycles in randomly augmented Dirac graphs—The complete collection*, *J. Graph Theory*. 2023;**104**:811–835. <https://doi.org/10.1002/jgt.23001>

### APPENDIX A: PROOF OF PROPOSITION 1.3

First we check by hand the statement for  $m \in \{7, 10, 11, \dots, 14\}$  (see Table A1).

As for  $m \geq 15$ , we will show that both

$$\lfloor \lambda_m \rfloor < (m - \lfloor \lambda_m \rfloor)(m - \lfloor \lambda_m \rfloor + 1) \quad (\text{A1})$$

and

$$\lceil \lambda_m \rceil < (m - \lceil \lambda_m \rceil)(m - \lceil \lambda_m \rceil + 1). \quad (\text{A2})$$

Since  $\lfloor \lambda_m \rfloor \leq \lambda_m$ , inequality (A1) will follow from

$$\lambda_m < (m - \lambda_m)(m - \lambda_m + 1).$$

Observe that, since  $\lambda_m = \sqrt{2m^2 + 2m}/2$ , we have

$$(m - \lambda_m)(m - \lambda_m + 1) - \lambda_m = \left( \frac{3m}{2} - \sqrt{2m(m+1)} \right) (m+1).$$

Thus, since  $\frac{3m}{2} - \sqrt{2m(m+1)} > 0$  for  $m \geq 9$ , the inequality (A1) holds.

Next, since  $\lceil \lambda_m \rceil < \lambda_m + 1$ , inequality (A2) will follow from

$$\lambda_m + 1 \leq (m - \lambda_m - 1)(m - \lambda_m).$$

**TABLE A1** Verifying inequality  $\ell_m < (m - \ell_m)(m - \ell_m + 1)$  for  $m \in \{7, 10, 11, \dots, 14\}$ .

$m$	7	10	11	12	13	14
$\lambda_m$	$2\sqrt{7}$	$\sqrt{55}$	$\sqrt{66}$	$\sqrt{78}$	$\sqrt{91}$	$\sqrt{105}$
$\lfloor \lambda_m \rfloor$	5	7	8	8	9	10
$\lceil \lambda_m \rceil$	6	8	9	9	10	11
$f_m(\lfloor \lambda_m \rfloor)$	$\frac{13}{5}$	$\frac{27}{7}$	$\frac{17}{4}$	$\frac{19}{4}$	$\frac{46}{9}$	$\frac{11}{2}$
$f_m(\lceil \lambda_m \rceil)$	$\frac{8}{3}$	$\frac{31}{8}$	$\frac{13}{3}$	$\frac{14}{3}$	$\frac{51}{10}$	$\frac{61}{11}$
$\ell_m$	5	7	8	9	10	10
$r_m = m - \ell_m$	2	3	3	3	3	4
$r_m(r_m + 1)$	6	12	12	12	12	20

Again, after some calculations, we get that

$$(m - \lambda_m - 1)(m - \lambda_m) - \lambda_m - 1 = \left(\frac{3m}{2} - \sqrt{2m(m + 1)} - \frac{1}{2}\right)m - 1.$$

Due to the arithmetic–geometric mean inequality we obtain that  $\sqrt{m(m + 1)} \leq (2m + 1)/2$ . Consequently,

$$\begin{aligned} \left(\frac{3m}{2} - \sqrt{2m(m + 1)} - \frac{1}{2}\right)m - 1 &\geq \left(\frac{3m}{2} - \sqrt{2} \cdot \frac{2m + 1}{2} - \frac{1}{2}\right)m - 1 \\ &= \left(\frac{3}{2} - \sqrt{2}\right)m^2 - \frac{\sqrt{2} + 1}{2}m - 1. \end{aligned}$$

One can easily check that the latter quadratic function of  $m$  is positive for  $m \geq 15$ . This completes the proof of (A2).

**APPENDIX B: PROOF OF CLAIM 4.5**

We can assume that  $r < \ell - 1$ , as the case  $\ell \in \{r, r + 1\}$  was already discussed at the beginning of the proof of Proposition 4.3. Recall that  $t \geq 2$ . Suppose that  $s = t$ , that is,  $V_i \subset H$  for  $i = 1, 2, \dots, t - 1$ , and set  $x = |V(H) \cap V_t|$ . Notice that the maximum number of edges with at least one endpoint in  $V(H) \cap V_t$  is achieved when all  $x$  vertices form the leftmost subset of  $V_t$ . Therefore, we assume this throughout [6].

Case  $0 \leq x \leq r$ : Our aim is to show that

$$d_t = \frac{t \binom{\ell}{2} + (t - 1) \binom{r+1}{2}}{t\ell - 1} > \frac{(t - 1) \binom{\ell}{2} + (t - 2) \binom{r+1}{2} + rx}{(t - 1)\ell + x - 1} = d_H,$$

where the second equality uses the fact that by our assumptions  $H$  consists of  $t - 1$  copies of  $K_\ell$ , such that each consecutive two of them are joined by an  $r$ -bridge and there is an additional  $x$ -tuple of vertices in  $V_t$  in which each vertex has exactly  $r$  neighbors to the left of it in  $H$ .

Consider the function

$$\begin{aligned} f(\ell, r, t, x) &= \left(t \binom{\ell}{2} + (t - 1) \binom{r + 1}{2}\right)((t - 1)\ell + x - 1) \\ &\quad - \left((t - 1) \binom{\ell}{2} + (t - 2) \binom{r + 1}{2} + rx\right)(t\ell - 1). \end{aligned}$$

If we show that for  $\ell, r, t, x$  satisfying our assumptions  $f(\ell, r, t, x) > 0$ , we will be done. Notice that  $f(\ell, r, t, x)$  is a linear function with respect to  $x$ , hence it is enough to verify that  $f(\ell, r, t, 0) > 0$  and  $f(\ell, r, t, r) > 0$ . We have

$$\begin{aligned} f(\ell, r, t, 0) &= \left(t \binom{\ell}{2} + (t - 1) \binom{r+1}{2}\right)(\ell t - \ell - 1) \\ &\quad - \left((t - 1) \binom{\ell}{2} + (t - 2) \binom{r+1}{2}\right)(\ell t - 1) \\ &= -\binom{\ell}{2} + \binom{r+1}{2}(\ell - 1) = \frac{\ell - 1}{2}(r(r + 1) - \ell) > 0, \end{aligned}$$

since  $r(r+1) > \ell$ . Similarly, since  $1-t < 0$  and  $1 \leq r < \ell-1$ , we get

$$\begin{aligned}
 f(\ell, r, t, r) &= \left( t \binom{\ell}{2} + (t-1) \binom{r+1}{2} \right) (\ell t - \ell + r - 1) \\
 &\quad - \left( (t-1) \binom{\ell}{2} + (t-2) \binom{r+1}{2} + r^2 \right) (\ell t - 1) \\
 &= \binom{\ell}{2} (rt - 1) + \binom{r+1}{2} (rt + \ell - r - 1) - \ell r^2 t + r^2 \\
 &= \frac{1}{2} (\ell^2 rt - \ell^2 - \ell rt + \ell + r^3 t + \ell r^2 - r^3 - r^2 + r^2 t + \ell r - r^2 - r - 2\ell r^2 t + 2r^2) \\
 &= \frac{1}{2} ((\ell r^2 - r^3) + (-\ell r^2 t + r^3 t) + (\ell^2 rt - \ell r^2 t) + (-\ell rt + r^2 t) + (-\ell^2 + \ell r) + (\ell - r)) \\
 &= \frac{\ell-r}{2} (r^2 - r^2 t + \ell rt - rt - \ell + 1) \\
 &= \frac{\ell-r}{2} (r^2(1-t) + (\ell-1)rt - \ell + 1) \\
 &> \frac{\ell-r}{2} ((\ell-1)r(1-t) + (\ell-1)rt - \ell + 1) \\
 &= \frac{\ell-r}{2} (\ell-1)(r-1) \geq 0.
 \end{aligned}$$

Case  $r+1 \leq x \leq \ell-1$ : This time we would like to show that

$$d_t = \frac{t \binom{\ell}{2} + (t-1) \binom{r+1}{2}}{t\ell - 1} > \frac{(t-1) \binom{\ell}{2} + (t-1) \binom{r+1}{2} + \binom{x}{2}}{(t-1)\ell + x - 1} = d_H,$$

where the second equality uses the fact that by our assumptions  $H$  consists of  $t-1$  copies of  $K_\ell$  and one copy of  $K_x$ , and since  $x \geq r+1$  each two consecutive cliques in  $H$  are joined by an  $r$  bridge.

Consider the function

$$\begin{aligned}
 g(\ell, r, t, x) &= \left( t \binom{\ell}{2} + (t-1) \binom{r+1}{2} \right) ((t-1)\ell + x - 1) \\
 &\quad - \left( (t-1) \binom{\ell}{2} + (t-1) \binom{r+1}{2} + \binom{x}{2} \right) (t\ell - 1).
 \end{aligned}$$

Once again we will show that  $g(\ell, r, t, x) > 0$  for  $\ell, r, t, x$  satisfying our assumptions. Before we proceed, notice that

$$\begin{aligned}
 g(\ell, r, t, x) &= \binom{\ell}{2} (tx - 1) + \binom{r+1}{2} (t-1)(x - \ell) - \binom{x}{2} (\ell t - 1) \\
 &= \frac{1}{2} ((\ell^2 tx - \ell^2 - \ell tx + \ell) \\
 &\quad + (r^2 tx - \ell r^2 t - r^2 x + \ell r^2 + rtx - \ell rt - rx + \ell r) \\
 &\quad + (-\ell tx^2 + x^2 + \ell tx - x)) \\
 &= \frac{1}{2} ((\ell^2 tx - \ell tx^2) + (-\ell r^2 t + r^2 tx) + (\ell r^2 - r^2 x) + (-\ell rt + rtx) \\
 &\quad + (-\ell^2 + \ell x) + (\ell r - rx) + (-\ell x + x^2) + (\ell - x)) \\
 &= \frac{\ell-x}{2} (\ell tx - r^2 t + r^2 - rt - \ell + r - x + 1).
 \end{aligned}$$



Hence, it is enough to show that

$$h(\ell, r, t, x) = \ell tx - r^2 t + r^2 - rt - \ell + r - x + 1 > 0.$$

As before, the function  $h$  is linear with respect to  $x$ . Furthermore, since  $\ell t - 1 > 0$  this function is increasing with respect to  $x$  and so  $h(\ell, r, t, x) \geq h(\ell, r, t, r + 1)$ . Thus, one only needs to check whether  $h(\ell, r, t, r + 1) > 0$ . Since  $r \leq \ell - 2$  and  $1 - t < 0$ , we have

$$\begin{aligned} h(\ell, r, t, r + 1) &= \ell t(r + 1) - r^2 t + r^2 - rt - \ell \\ &= (\ell - 1)rt + r^2(1 - t) + \ell(t - 1) \\ &\geq (\ell - 1)rt + (\ell - 2)r(1 - t) + \ell(t - 1) \\ &= (r + t - 1)\ell + r(t - 2) > 0. \end{aligned}$$

This completes the proof of Proposition 4.3. □