European Journal of Combinatorics 91 (2021) 103200

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

An analogue of the Erdős–Gallai theorem for random graphs

József Balogh [a,b,1], Andrzej Dudek [c,2], Lina Li [d]

[a] *Department of Mathematical Sciences, University of Illinois at Urbana-Champaign, IL, USA*
[b] *Moscow Institute of Physics and Technology, Russian Federation*
[c] *Department of Mathematics, Western Michigan University, Kalamazoo, MI, USA*
[d] *Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

ARTICLE INFO

Article history:
Available online 25 August 2020

ABSTRACT

Recently, variants of many classical extremal theorems have been proved in the random environment. We, complementing existing results, extend the Erdős–Gallai Theorem in random graphs. In particular, we determine, up to a constant factor, the maximum number of edges in a P_n-free subgraph of $G(N, p)$, practically for all values of N, n and p. Our work is also motivated by the recent progress on the size-Ramsey number of paths.

1. Introduction

A celebrated theorem of Erdős and Gallai [14] from 1959 determines the maximum number of edges in an n-vertex graph with no k-vertex path P_k.

Theorem 1 (*Erdős and Gallai [14]*)**.** *For $n, k \geq 2$, if G is an n-vertex graph with no copy of P_k, then the number of edges of G satisfies $e(G) \leq \frac{1}{2}(k-2)n$. If n is divisible by $k-1$, then the maximum is achieved by a union of disjoint copies of K_{k-1}.*

An important direction of combinatorics in recent years is the study of sparse random analogues of classical extremal results; that is, the extent to which of these results remain true in a random setting. For graphs G and F, we write $\mathrm{ex}(G, F)$ for the maximum number of edges in an F-free

E-mail addresses: jobal@illinois.edu (J. Balogh), andrzej.dudek@wmich.edu (A. Dudek), linali2@illinois.edu (L. Li).

1 Partially supported by NSF, USA Grant DMS-1764123 and Arnold O. Beckman Research Award (UIUC) Campus Research Board 18132, USA and the Langan Scholar Fund (UIUC), USA.

2 Partially supported by Simons Foundation, USA Grant #522400.

https://doi.org/10.1016/j.ejc.2020.103200
0195-6698/

subgraph of G. For example, the Erdős–Gallai theorem asserts that $\mathrm{ex}(K_n, P_k) = \frac{1}{2}(k-2)n$ if n is divisible by $k-1$.

The study of the random variable $\mathrm{ex}(G, F)$, where G is the Erdős–Rényi random graph $G(n, p)$, was initiated by Babai, Simonovits and Spencer [2], and by Frankl and Rödl [15]. After efforts by several researchers [18,19,21–23,32], Conlon and Gowers [9] and Schacht [30] finally proved a sparse random version of the Erdős–Stone theorem, showing a *transference principle* of Turán-type results, that is, when a random graph inherits its (relative) extremal properties from the classical deterministic case. Note that via the hypergraph container method the same results were proved [4] and [29], even when $|F|$ is a reasonable large function of n. A special case of this result, when F is the k-vertex path P_k, can be viewed as a weak analogue (as the Turán density is 0) of the Erdős–Gallai theorem on the random graph for paths with a fixed size. In this paper, we investigate the random analogue of the Erdős–Gallai theorem for general paths, whose length might increase with the order of the random graph.

We say that events A_n in a probability space hold *asymptotically almost surely* (or a.a.s.), if the probability that A_n holds tends to 1 as n goes to infinity. The typical appearance of long paths and cycles is one of the most thoroughly studied direction in random graph theory. Over the past decades, there were many diverse and beautiful results on this subject. In a seminal paper, Ajtai, Komlós and Szemerédi [1], confirming a conjecture of Erdős, proved that for $p = \frac{c}{n}$ with $c > 1$, $G(n, p)$ contains a path of length $\alpha(c)n$ a.a.s. where $\lim_{c\to\infty} \alpha(c) = 1$. Frieze [16] later determined the asymptotics of the number of vertices not covered by a longest path in $G(n, p)$. For Hamiltonicity, Bollobás [8] and Komlós and Szemerédi [24] independently proved that for $p \geq \frac{\log n+\log\log n+\omega(1)}{n}$, the random graph $G(n, p)$ is a.a.s. Hamiltonian. Turán-type results for long cycles in $G(n, p)$ was also studied under the name of *global resilience*, that is, the minimum number r such that one can destroy the graph property by deleting r edges. Dellamonica Jr, Kohayakawa, Marciniszyn and Steger [10] determined the global resilience of $G(n, p)$ with respect to the property of containing a cycle of length proportional to the number of vertices. Very recently, Krivelevich, Kronenberg and Mond [27] studied the transference principle in the context of long cycles and in particular showing the following.

Theorem 2 (*Corollary 1.10 in [27]*). *For every $0 < \beta < \frac{1}{5}$, there exists $C > 0$ such that if $G = G(N, p)$ where $p \geq \frac{C}{N}$, then for any $\frac{C_1}{\log(1/\beta)} \cdot \log N \leq n \leq (1 - C_2\beta)N$, with probability $1 - e^{\Omega(N)}$,*

$$\mathrm{ex}(G(N,p), C_n) \leq \left(\frac{\mathrm{ex}(K_N, C_n)}{\binom{N}{2}} + \beta \right) e(G(N,p)), \tag{1}$$

where $C_1, C_2 > 0$ are absolute constants.

We aim to explore the global resilience of general long paths. More formally, given integers $N > n$, we are interested in determining the asymptotic behavior of random variable $\mathrm{ex}(G(N, p), P_{n+1})$ as N and n go to infinity at the same time.

We start with an observation, which is proved in Section 3.

Proposition 3. *For every $\frac{1}{N^2} \ll p \leq \frac{1}{N}$ and $n \geq 2$, a.a.s. we have $\mathrm{ex}(G(N, p), P_{n+1}) = \Theta(pN^2)$. In particular, a.a.s. $\mathrm{ex}(G(N, 1/N), P_{n+1}) \geq N/15$.*

Therefore, throughout this paper, we naturally restrict ourselves to the regime $p \geq 1/N$ and have the following trivial lower bound

$$a.a.s. \quad \mathrm{ex}(G(N,p), P_{n+1}) \geq \mathrm{ex}\left(G\left(N, 1/N\right), P_{n+1}\right) \geq N/15. \tag{2}$$

We prove the following results.

Theorem 4. *Let $3n \leq N \leq ne^{2n}$. The following hold a.a.s. as n approaches infinity.*

(i) *For $p \geq \left(\log \frac{N}{n}\right)/(6n)$, we have $\frac{1}{4}pnN \leq \mathrm{ex}(G(N, p), P_{n+1}) \leq 18pnN$.*

(ii) *Let* $\omega = \left(\log \frac{N}{n}\right)/(np)$*. For* $N^{-1} \le p \le \left(\log \frac{N}{n}\right)/(6n)$*, we have*

$$\frac{1}{75}\frac{\omega}{\log\omega}pnN \le \mathrm{ex}(G(N,p), P_{n+1}) \le 8\frac{\omega}{\log\omega}pnN.$$

Theorem 5. *Let* $N \ge ne^{2n}$*. The following hold a.a.s. as n approaches infinity.*

(i) *For* $p \ge N^{-\frac{2}{5n}}$*, we have* $\frac{1}{16}nN \le \mathrm{ex}(G(N,p), P_{n+1}) \le \frac{1}{2}nN$*.*
(ii) *Let* $\omega = (\log N)/(np)$*. For* $N^{-1} \le p \le N^{-\frac{2}{5n}}$*, we have*

$$\frac{1}{75}\frac{\omega}{\log\omega}pnN \le \mathrm{ex}(G(N,p), P_{n+1}) \le 8\frac{\omega}{\log\omega}pnN.$$

Remark 1. Assume that n is even. Then (1) together with $\mathrm{ex}(K_N, C_n) \le nN^{1+2/n}$ [28] implies that

$$\begin{aligned}\mathrm{ex}(G(N,p), P_n) \le \mathrm{ex}(G(N,p), C_n) &\le \left(\frac{\mathrm{ex}(K_N, C_n)}{\binom{N}{2}} + \beta\right) e(G(N,p)) \\ &\le \left(\frac{nN^{1+2/n}}{\binom{N}{2}} + \beta\right)\frac{pN^2}{2} \sim pnN^{1+2/n} + \beta\frac{pN^2}{2},\end{aligned}$$

which is weaker than our bounds. (Recall that $p \ge \frac{C}{n}$, where $C = C(\beta)$.) Of course, there are some better upper bounds for $\mathrm{ex}(K_N, C_n)$, which could be used to make an improvement. However, since, in general, $\mathrm{ex}(K_N, C_n)$ behaves differently with $\mathrm{ex}(K_N, P_n)$ and is indeed much greater, Krivelevich, Kronenberg, and Mond's result [27] and ours do not imply one another.

Remark 2. One can run the same proof and show that Theorem 5 holds when n is a constant greater than 1 and N approaches infinity. Note also that a result of Johansson, Kahn and Vu [20] on the threshold function of the property that $G(N,p)$ contains a K_n-factor (n is a constant) implies $\mathrm{ex}(G(N,p), P_{n+1}) = \frac{1}{2}(n-1)N$ for $p = \Omega\left(N^{-2/n}(\log n)^{1/\binom{n}{2}}\right)$, whenever N is divisible by n. Indeed, they determined the threshold function for containing a H-factor (H is a fixed graph), which might be useful for further improving the above result.

We made no attempt to optimize the constants in the theorems. Throughout the paper, we omit all floor and ceiling signs whenever these are not crucial. All logarithms in this paper have base e.

2. Tools

In this section, we list several results that we will use. The first lemma is a direct application of the depth first search algorithm (DFS), which has appeared in [12]. Using the DFS algorithm in finding long paths was first introduced by Ben-Eliezer, Krivelevich, and Sudakov [6], and then it became a particularly suitable tool in this topic.

Lemma 6 ([12])**.** *For every* P_{n+1}*-free graph H on N vertices, we can find a decomposition of edges into* $\bigcup_{i=1}^{N/n} F_i$*, where* $F_i = E(S_i) \cup E(S_i, T_i)$ *for two disjoint sets* $S_i, T_i \subseteq [N]$ *with* $|S_i| = |T_i| = n$*.*

We also need the following form of Chernoff's bound.

Lemma 7 (*Chernoff's Bound*)**.** *Let* $X = \sum_{i=1}^{n} X_i$*, where* $X_i = 1$ *with probability* p_i *and* $X_i = 0$ *with probability* $1 - p_i$*, and all* X_i*'s are independent. Let* $\mu = \mathbb{E}(X) = \sum_{i=1}^{n} p_i$*. Then, for all* $0 < \delta < 1$*,*

$$\mathbb{P}(X \le (1-\delta)\mu) \le e^{-\mu\delta^2/2}.$$

The third lemma is a key ingredient of our proof, which is used to find dense subsets in random graphs. This may be of independent interest.

Lemma 8. *For $N > 2n$, $0 < p < 1$ and a constant $0 < \alpha \le 1/2$, let $r = N/n$ and choose an arbitrary β satisfying*

$$\max\left\{2\log(2e),\ \frac{2}{\alpha np}\log\left(\frac{1}{\alpha np}\right)\right\} \le 2\beta\log\beta \le \min\left\{2\left(\frac{1}{p}\right)\log\left(\frac{1}{p}\right),\ \frac{1}{np}\left(\log r - \log\alpha 2^{\frac{1}{\alpha}}\right)\right\}. \tag{3}$$

Then there exists a positive constant $c = c(\alpha)$ such that with probability at least $1 - \exp(-cr^{\alpha}n)$ there exists an n-set in $G(N,p)$ with at least $\left(\frac{1-\alpha}{2}\right)\beta pn^2$ edges.

Remark 3. Lemma 8 essentially states that given N, n, for some range of p, we can find an n-vertex subgraph, which is denser than the random graph by some factor β. For instance, as it will be explained in the proof of Theorem 4(ii), when $135n \le N \le ne^{2n}$, we can choose $\frac{\log r}{nr^{1/5}} \le p \le \frac{\log r}{6n}$, so that $2\beta\log\beta = \frac{1}{np}\log\left(\frac{3}{8}r\right)$ satisfying (3). Note that if $p \ll \frac{\log r}{n}$, we have $\beta = \omega(1)$, and therefore the graph we produce here is much denser than the random graph.

Proof. One can check that the function $f(x) = x\log x$ is non-negative and increasing for $x \ge 1$. Thus, $\log(2e) \le f(\beta) \le f(1/p)$ implies that

$$\max\left\{2, \frac{1}{\alpha np}\right\} < \beta \le 1/p. \tag{4}$$

Let $B_0 = [N]$. We will construct the desired set iteratively. In each step, take an arbitrary subset $A_i \subseteq B_{i-1}$ of size αn, and let

$$B_i = \{v \in B_{i-1} \setminus A_i :\ \deg(v, A_i) \ge \beta\alpha np\}.$$

We will show that a.a.s. we can continue this process in $\lceil \frac{1}{\alpha} \rceil$ steps. For convenience, in the rest of the proof, we ignore all floor and ceiling signs.

Claim 9. *$|B_i| \ge \frac{rn}{2^i}\exp(-2i\beta\log\beta\cdot\alpha np)$, for all $0 \le i \le \frac{1}{\alpha} - 1$ with probability at least $1 - \exp(-\Omega(r^{\alpha}n))$.*

We prove it by induction on $i \ge 0$. For $i = 0$, it is trivial. Suppose the statement holds for $i - 1$. That means

$$|B_{i-1}| \ge \frac{rn}{2^{i-1}}\exp\left(-2(i-1)\beta\log\beta\cdot\alpha np\right) \tag{5}$$

with probability at least $1 - \exp(-\Omega(r^{\alpha}n))$. Furthermore, $0 \le i \le \frac{1}{\alpha} - 1$ yields that $(i-1)\alpha < i\alpha \le 1 - \alpha < 1$ and hence,

$$|B_{i-1}| \ge \frac{rn}{2^{\frac{1}{\alpha}-2}}\exp\left(-2\beta\log\beta\cdot np\right) \ge \frac{rn}{2^{\frac{1}{\alpha}-2}}\exp\left(-\left(\log r - \log\alpha 2^{\frac{1}{\alpha}}\right)\right) = 4\alpha n,$$

consequently

$$|B_{i-1}| - \alpha n \ge \frac{3}{4}|B_{i-1}| > \frac{|B_{i-1}|}{\sqrt{2}}.$$

Then, the expected size of B_i is

$$\mathbb{E}(|B_i|) = (|B_{i-1}| - \alpha n)\mathbb{P}(\deg(v, A_i) \ge \beta\alpha np) \ge \frac{1}{\sqrt{2}}|B_{i-1}|\binom{\alpha n}{\beta\alpha np}p^{\beta\alpha np}(1-p)^{\alpha n}.$$

Due to (4), we get that $p \le 1/\beta \le 1/2$ and $\beta\alpha np \ge 1$. Now we use $\binom{\alpha n}{\beta\alpha np} \ge \left(\frac{\alpha n}{\beta\alpha np}\right)^{\beta\alpha np} = \left(\frac{1}{\beta p}\right)^{\beta\alpha np}$ and the inequality $1-p \ge (2e)^{-p}$, which is valid for $0 \le p \le 1/2$. Thus,

$$\begin{aligned}\mathbb{E}(|B_i|) &= (|B_{i-1}| - \alpha n)\mathbb{P}(\deg(v, A_i) \ge \beta\alpha np) \ge \frac{1}{\sqrt{2}}|B_{i-1}| \exp(-(\beta \log \beta + \log 2e)\alpha np)\\ &\ge \frac{1}{\sqrt{2}}|B_{i-1}| \exp(-2\beta \log \beta \cdot \alpha np).\end{aligned}$$

Observe that conditioning on (5) gives

$$\begin{aligned}\mathbb{E}(|B_i|) &\ge \frac{1}{\sqrt{2}}|B_{i-1}| \exp(-2\beta \log \beta \cdot \alpha np)\\ &\ge \frac{1}{\sqrt{2}} \cdot \frac{rn}{2^{i-1}} \exp\left(-2(i-1)\beta \log \beta \cdot \alpha np\right) \cdot \exp(-2\beta \log \beta \cdot \alpha np)\\ &= \frac{1}{\sqrt{2}} \cdot \frac{rn}{2^{i-1}} \exp\left(-2i\beta \log \beta \cdot \alpha np\right) \ge \frac{1}{\sqrt{2}} \cdot \frac{rn}{2^{i-1}} \exp\left(-\alpha i \left(\log r - \log \alpha 2^{\frac{1}{\alpha}}\right)\right)\\ &\ge \frac{1}{\sqrt{2}} \cdot \frac{rn}{2^{\frac{1}{\alpha}-1}} \exp\left(-(1-\alpha)\left(\log r - \log \alpha 2^{\frac{1}{\alpha}}\right)\right) = \Omega(r^\alpha n),\end{aligned}$$

which goes to infinity together with n. Therefore, Chernoff's bound (applied with $\delta = 1 - 1/\sqrt{2}$) yields that with probability at least $1 - \exp(-\Omega(r^\alpha n))$ we have

$$|B_i| \ge \frac{1}{\sqrt{2}}\mathbb{E}(|B_i|) \ge \frac{1}{2}|B_{i-1}| \exp(-2\beta \log \beta \cdot \alpha np) \ge \frac{rn}{2^i} \exp\left(-2i\beta \log \beta \cdot \alpha np\right),$$

where the last inequality follows from (5).

Now we finish the proof of Lemma 8. Claim 9 gives that with probability at least $1 - \exp(-\Omega(r^\alpha n))$ the set $B_{\frac{1}{\alpha}-1}$ exists and satisfies

$$\left|B_{\frac{1}{\alpha}-1}\right| \ge \frac{rn}{2^{\frac{1}{\alpha}-1}} \exp\left(-\left(\log r - \log \alpha 2^{\frac{1}{\alpha}}\right)\right) = 2\alpha n > \alpha n.$$

Therefore, we can find disjoint sets $A_1, \dots, A_{1/\alpha}$ of size αn with $e(A_i, A_j) \ge \alpha n \cdot \beta\alpha np$ for all $1 \le i < j \le 1/\alpha$. Let $A = \bigcup_{i=1}^{1/\alpha} A_i$. Then we have $|A| = n$ and

$$e(A) \ge \binom{1/\alpha}{2}\alpha n \cdot \beta\alpha np = \binom{1-\alpha}{2}\beta pn^2. \quad \square$$

We also present the following two probabilistic results which will be used later.

Lemma 10. *Assume that $np \ge \left(\log \frac{N}{n}\right)/6$ and $N \ge 3n$. Then a.a.s. for every two disjoint sets $S, T \subseteq [N]$, $|S| = |T| = n$, the number of edges in $G \in G(N,p)$ induced by $S \cup T$ with at least one endpoint in S is at most $18n^2p$.*

Proof. Let $X_{S,T}$ be the number of edges in $G(N,p)$ with one endpoint in S and one endpoint in T. Observe that $\mathbb{E}(X_{S,T}) = \left(\frac{3}{2} - \frac{1}{2n}\right)n^2p$. Note that if $3n^2/2 \le 18n^2p$, then the statement is trivial. Otherwise, the union bound implies that

$$\begin{aligned}\mathbb{P}(\exists S, T, X_{S,T} \ge 18n^2p) &\le \binom{N}{n}^2 \binom{3n^2/2}{18n^2p} p^{18n^2p} \le \left(\frac{Ne}{n}\right)^{2n} \left(\frac{e}{12}\right)^{18n^2p}\\ &= \exp\left(-n\left(18np \log\left(\frac{12}{e}\right) - 2\log\left(\frac{Ne}{n}\right)\right)\right).\end{aligned}$$

Since $np \geq \left(\log \frac{N}{n}\right)/6$ and $N \geq 3n$, we obtain that

$$18np \log\left(\frac{12}{e}\right) - 2\log\left(\frac{Ne}{n}\right) \geq 3\log\left(\frac{12}{e}\right)\log\left(\frac{N}{n}\right) - 2\log\left(\frac{Ne}{n}\right)$$
$$\geq 4\log\left(\frac{N}{n}\right) - 2\log\left(\frac{N}{n}\right) - 2 = 2\log\left(\frac{N}{n}\right) - 2 \geq 2\log 3 - 2 \geq 0.19.$$

Finally, we conclude that $\mathbb{P}(\exists S, T, X_{S,T} \geq 18n^2p) \leq \exp(-0.19n) = o(1)$, which completes the proof. □

Lemma 11. *Let* $\beta = \frac{\frac{1}{np}\log\frac{N}{n}}{\log\left(\frac{1}{np}\log\frac{N}{n}\right)} > 1$ *and* $m = 8\beta n^2 p$. *Then a.a.s. for every two disjoint sets* $S, T \subseteq [N]$, $|S| = |T| = n$, *the number of edges induced by* $S \cup T$ *with at least one endpoint in* S *is at most* m.

Proof. We assume $m < 3n^2/2$ since otherwise Lemma 11 holds trivially. By a simple union bound, we obtain

$$\mathbb{P}(\exists S, T, X_{S,T} \geq m) \leq \binom{N}{n}^2 \binom{3n^2/2}{m} p^m \leq \exp\left(2n\log\left(\frac{Ne}{n}\right)\right)\exp(-\log\beta \cdot m)$$
$$= \exp\left(2n\log\left(\frac{Ne}{n}\right) - 8\beta\log\beta \cdot n^2p\right).$$

Now we bound from below $\beta\log\beta$ by

$$\beta\log\beta = \frac{\frac{1}{np}\log\frac{N}{n}}{\log\left(\frac{1}{np}\log\frac{N}{n}\right)}\log\left(\frac{\frac{1}{np}\log\frac{N}{n}}{\log\left(\frac{1}{np}\log\frac{N}{n}\right)}\right) \geq \frac{\frac{1}{np}\log\frac{N}{n}}{\log\left(\frac{1}{np}\log\frac{N}{n}\right)}\log\sqrt{\frac{1}{np}\log\frac{N}{n}} = \frac{1}{2np}\log\left(\frac{N}{n}\right).$$

Thus,

$$\mathbb{P}(\exists S, T, X_{S,T} \geq m) \leq \exp\left(2n\log\left(\frac{Ne}{n}\right) - 8\beta\log\beta \cdot n^2p\right)$$
$$\leq \exp\left(2n\log\left(\frac{Ne}{n}\right) - \frac{4}{np}\log\left(\frac{N}{n}\right) \cdot n^2p\right)$$
$$\leq \exp\left(-n\left(4\log\left(\frac{N}{n}\right) - 2\log\left(\frac{Ne}{n}\right)\right)\right) = o(1),$$

where the last inequality follows from $N \geq 3n$ as $4\log\left(\frac{N}{n}\right) - 2\log\left(\frac{Ne}{n}\right) = 2\log\left(\frac{N}{n}\right) - 2 \geq 2\log(3) - 2 \geq 0.19$. □

3. Proofs of the main results

3.1. *Proof of Proposition 3*

Let $G = (V, E) = G(N, p)$. We will count the number of isolated edges. For a given pair of vertices $e \in \binom{V}{2}$, let X_e be an indicator random variable that takes value 1 if e is an isolated edge in G. Set $X = \sum_e X_e$. Observe that $\Pr(X_e = 1) = p(1-p)^{2(N-2)}$ and so

$$\mathbb{E}(X) = \binom{N}{2}p(1-p)^{2(N-2)} \sim \binom{N}{2}pe^{-2pN} \geq \binom{N}{2}pe^{-2} \to \infty,$$

by assumption. Furthermore, since for vertex disjoint $e, f \in \binom{V}{2}$, $\Pr(X_e = X_f = 1) = p^2(1-p)^{4(n-4)+4}$, we obtain that

$$\mathbb{E}(X^2) = \mathbb{E}(X) + \sum_{e \cap f = \emptyset} \Pr(X_e = X_f = 1) = \mathbb{E}(X) + 6\binom{N}{4}p^2(1-p)^{4(N-4)+4}.$$

Thus,

$$\frac{\mathbb{E}(X^2)}{\mathbb{E}(X)^2} = \frac{1}{\mathbb{E}(X)} + \frac{(N-2)(N-3)}{N(N-1)(1-p)^4} \le \frac{1}{\mathbb{E}(X)} + \frac{1}{(1-p)^4} \le \frac{1}{\mathbb{E}(X)} + \frac{1}{1-4p}$$

and

$$\frac{\mathrm{Var}(X)}{\mathbb{E}(X^2)} \le \frac{1}{\mathbb{E}(X)} + \frac{1}{1-4p} - 1 = \frac{1}{\mathbb{E}(X)} + \frac{4p}{1-4p} = o(1),$$

since $\mathbb{E}(X) \to \infty$ and also by assumption $p \to 0$. Now Chebyshev's inequality yields that X is concentrated around its mean and consequently a.a.s. we have

$$\mathrm{ex}(G(N,p), P_{n+1}) \ge (1+o(1))\mathbb{E}(X) = \Omega(pN^2).$$

The upper bound easily follows from the fact that $\mathrm{ex}(G(N,p), P_{n+1}) \le e(G(N,p))$.

Finally observe that a.a.s.

$$\mathrm{ex}(G(N,1/N), P_{n+1}) \ge (1+o(1))\mathbb{E}(X) \ge (1+o(1))\binom{N}{2}\frac{1}{N}e^{-2} \ge N/15. \quad \square$$

3.2. *Proof of Theorem 4*

Proof of Theorem 4(i). This proof is by now quite standard which applies the DFS algorithm and the first moment method. Recall that $np \ge \left(\log \frac{N}{n}\right)/6$ and $N \ge 3n$.

Observe that Lemma 10 together with Lemma 6 implies that for every P_{n+1}-free subgraph H of $G \in G(N,p)$ a.a.s.

$$e(H) \le \frac{N}{n} \cdot 18n^2 p = 18pnN,$$

which establishes the upper bound.

For the lower bound, take an arbitrary vertex partition $[N] = \bigcup_{i=1}^{N/n} S_i$, where $|S_i| = n$ for all i. Let H be the subgraph of $G \in G(N,p)$ whose edge set is $\bigcup E(G[S_i])$. Clearly, H is P_{n+1}-free. Note that $\mathbb{E}(e(H)) = \frac{N}{n}\left(\frac{1}{2} - \frac{1}{2n}\right)n^2 p = \left(\frac{1}{2} - \frac{1}{2n}\right)pnN$. By Chernoff's bound,

$$\mathbb{P}\left(e(H) \le \frac{1}{4}pnN\right) \le \exp\left(-\Omega(pnN)\right) = o(1),$$

since $pnN \to \infty$. Therefore, a.a.s. we have $\mathrm{ex}(G(N,p), P_{n+1}) \ge e(H) \ge \frac{1}{4}pnN$. $\square$

Proof of Theorem 4(ii). We first show the upper bound. Let $\beta_1 = \frac{\frac{1}{np}\log\frac{N}{n}}{\log\left(\frac{1}{np}\log\frac{N}{n}\right)}$ and $m = 8\beta_1 n^2 p$. Since $np \le \left(\log \frac{N}{n}\right)/6$, we know that $\beta_1 > 1$.

For every P_{n+1}-free subgraph H of $G \in G(N,p)$, Lemmas 6 and 11 imply that a.a.s

$$e(H) \le \frac{N}{n} \cdot m = 8\beta_1 pnN = 8\frac{\frac{1}{np}\log\frac{N}{n}}{\log\left(\frac{1}{np}\log\frac{N}{n}\right)}pnN,$$

which establishes the upper bound.

For the lower bound, we shall divide the discussion into three cases. First, let us assume $N \le 135n$. Together with $\frac{1}{np}\log\left(\frac{N}{n}\right) \ge 6 \ge e$, we have

$$\frac{\omega}{\log\omega}pnN = \frac{\log\left(\frac{N}{n}\right)}{\log\left(\frac{1}{np}\log\left(\frac{N}{n}\right)\right)}N \le \log\left(\frac{N}{n}\right)N < 5N.$$

Therefore, by (2), we trivially have

$$\mathrm{ex}(G(N,p), P_{n+1}) \ge N/15 \ge \frac{1}{75}\frac{\omega}{\log\omega}pnN.$$

Next, let us assume $p \le \log\left(\frac{N}{n}\right) / \left(n\left(\frac{N}{n}\right)^{1/5}\right)$. Similarly, we complete the proof by observing that

$$\frac{\omega}{\log\omega} pnN = \frac{\log\left(\frac{N}{n}\right)}{\log\left(\frac{1}{np}\log\left(\frac{N}{n}\right)\right)} N \le \frac{\log\left(\frac{N}{n}\right)}{\frac{1}{5}\log\left(\frac{N}{n}\right)} N = 5N.$$

It remains to prove the lower bound for the case when $N \ge 135n$ and

$$\frac{\log\left(\frac{N}{n}\right)}{n\left(\frac{N}{n}\right)^{1/5}} \le p \le \frac{\log(\frac{N}{n})}{6n}. \tag{6}$$

Indeed, such range of p only exists for $N \ge 6^5 n$. In this case, we will apply Lemma 8 repeatedly to find a dense subgraph with no P_{n+1}. Let

$$2\beta_2 \log\beta_2 = \min\left\{2\left(\frac{1}{p}\right)\log\left(\frac{1}{p}\right),\ \frac{1}{np}\log\left(\frac{3N}{8n}\right)\right\}.$$

Since $N \le ne^{2n}$ and $p \le \log\left(\frac{N}{n}\right)/(6n) \le \frac{1}{3}$, we have

$$2\left(\frac{1}{p}\right)\log\left(\frac{1}{p}\right) \ge 2\left(\frac{1}{p}\right)\log 3 > \frac{2}{p} \ge \frac{1}{np}\log\left(\frac{3N}{8n}\right).$$

Furthermore, since $N \ge 6^5 n$, we obtain

$$\log\left(\frac{3N}{8n}\right) \ge \log\left(\frac{3}{8}\right) + \frac{1}{5}\log 6^5 + \frac{4}{5}\log\left(\frac{N}{n}\right) > \frac{4}{5}\log\left(\frac{N}{n}\right),$$

and

$$2\beta_2\log\beta_2 = \frac{1}{np}\log\left(\frac{3N}{8n}\right) \ge \frac{4}{5np}\log\left(\frac{N}{n}\right) > 2\log(2e).$$

Finally, observe that for $\alpha = 1/2$,

$$\frac{1}{np}\log\left(\frac{3N}{8n}\right) \ge \frac{1}{np}\cdot 4\log\left(\frac{2\left(\frac{N}{n}\right)^{1/5}}{\log\left(\frac{N}{n}\right)}\right) \ge \frac{2}{\alpha np}\log\left(\frac{1}{\alpha np}\right),$$

where the first inequality is given by $N \ge 135n$ and the last inequality follows from (6). Thus, we can iteratively apply Lemma 8 $N/4n$ times with $\alpha = \frac{1}{2}$ and $r = \frac{3N}{4n}$ and find $N/4n$ disjoint n-sets A_i, where a.a.s. for all i

$$e(A_i) \ge \left(\frac{1-\alpha}{2}\right)\beta_2 pn^2 \ge \frac{1-\alpha}{4}\frac{\frac{1}{np}\log\left(\frac{3N}{8n}\right)}{\log\left(\frac{1}{np}\log\left(\frac{3N}{8n}\right)\right)} pn^2 \ge \frac{1}{10}\frac{\frac{1}{np}\log\left(\frac{N}{n}\right)}{\log\left(\frac{1}{np}\log\left(\frac{N}{n}\right)\right)} pn^2.$$

Let H be the subgraph of G with vertex set $\bigcup_{i=1}^{N/4n} A_i$, and edge set $\bigcup_{i=1}^{N/4n} E(A_i)$. Note that H is P_{n+1}-free and therefore, a.a.s. we have

$$\text{ex}(G(N,p), P_{n+1}) \ge e(H) \ge \frac{1}{10}\frac{\frac{1}{np}\log\left(\frac{N}{n}\right)}{\log\left(\frac{1}{np}\log\left(\frac{N}{n}\right)\right)} pn^2 \cdot \frac{N}{4n} = \frac{1}{40}\frac{\frac{1}{np}\log\left(\frac{N}{n}\right)}{\log\left(\frac{1}{np}\log\left(\frac{N}{n}\right)\right)} pnN. \quad \square$$

3.3. *Proof of Theorem 5*

Proof of Theorem 5(i). By the Erdős–Gallai Theorem (Theorem 1), it is sufficient to prove the lower bound. Let

$$2\beta\log\beta = \min\left\{2\left(\frac{1}{p}\right)\log\left(\frac{1}{p}\right),\ \frac{4}{5np}\log N\right\}.$$

Since $p \geq N^{-\frac{2}{5n}}$, we have $\beta = 1/p$. If $p > 1/3$, then the proof simply follows from the proof of Theorem 4(i). Otherwise, we have $2\beta \log \beta \geq 6 \log 3 > 2 \log(2e)$. Similarly as in the proof of Theorem 4(ii), we can iteratively apply Lemma 8 $N/4n$ times with $\alpha = \frac{1}{2}$ and $r = \frac{3N}{4n}$, and a.a.s. find a P_{n+1}-free subgraph H of $G(N, p)$ with

$$e(H) \geq \left(\frac{1-\alpha}{2}\right) \beta p n^2 \cdot \frac{N}{4n} = \frac{1}{16} nN. \quad \square$$

Proof of Theorem 5(ii). The proof of the upper bound is the same as in Theorem 4(ii) and we skip here the full details. For the lower bound, we first assume that $p < N^{-1/5}$. Observe that

$$\frac{\omega}{\log \omega} pnN = \frac{\log N}{\log\left(\frac{1}{np}\log N\right)} N \leq \frac{\log N}{\log N^{1/5}} N = 5N,$$

where the inequality holds for $N \geq ne^{2n}$. Therefore, by (2), we trivially have

$$\mathrm{ex}(G(N,p), P_{n+1}) \geq N/15 \geq \frac{1}{75} \frac{\omega}{\log \omega} pnN.$$

It remains to show the lower bound for $p \geq N^{-1/5}$. Let

$$2\beta \log \beta = \min\left\{2\left(\frac{1}{p}\right) \log\left(\frac{1}{p}\right),\ \frac{4}{5np} \log N\right\}.$$

Since $p \leq N^{-\frac{2}{5n}}$, we have $2\beta \log \beta = \frac{4}{5np} \log N$. Since $N \geq ne^{2n}$, we have

$$\frac{1}{np} \log\left(\frac{3N}{8n}\right) \geq 2\beta \log \beta \geq \frac{4}{5np} \log\left(ne^{2n}\right) \geq \frac{8}{5p} \geq \frac{8}{5} N^{\frac{2}{5n}} \geq \frac{8e^{\frac{4}{5}}}{5} > 2\log(2e).$$

Moreover, observe that for $\alpha = \frac{1}{2}$ and $p \geq N^{-1/5}$, we have $2\beta \log \beta \geq \frac{2}{\alpha np} \log\left(\frac{1}{\alpha np}\right)$. Similarly as in the proof of Theorem 4(ii), the proof is completed by iteratively applying Lemma 8 $N/4n$ times with $\alpha = \frac{1}{2}$ and $r = \frac{3N}{4n}$. $\square$

4. Long paths and multicolor size-Ramsey number

The size-Ramsey number $\hat{R}(F, r)$ of a graph F is the smallest integer m such that there exists a graph G on m edges with the property that any r-coloring of the edges of G yields a monochromatic copy of F. The study of size-Ramsey number was initiated by Erdős, Faudree, Rousseau and Schelp [13]. For paths, Beck [5], resolving a \$100 question of Erdős, proved that $\hat{R}(P_n, 2) < 900n$ for sufficiently large n. The strongest upper bound, $\hat{R}(P_n, 2) \leq 74n$, was given by Dudek and Prałat [11], and they also provide the lower bound, $\hat{R}(P_n, 2) \geq 5n/2 - O(1)$. Very recently, Bal and DeBiasio [3] further improved the lower bound to $(3.75 - o(1))n$.

For more colors, it was proved in [11] that $\frac{(r+3)r}{4} n - O(r^2) \leq \hat{R}(P_n, r) \leq 33r4^r n$. Subsequently, Krivelevich [26] (see also [25]) showed that $\hat{R}(P_n, r) = O((\log r) r^2 n)$. An alternative proof of the above result was later given by Dudek and Prałat [12]. Both proofs indeed give a stronger *density-type* result, which shows that any dense subset of a large enough structure contain the desired substructure. In particular, the proof in [12] implies the following result.

Theorem 12 *([12]). For $r \geq 2$ and $c \geq 7$, there exists a constant $\alpha = \alpha(c)$ such that the following statement holds a.a.s. for $p \geq \alpha(\log r)/n$. Every subgraph H of $G \in G(crn, p)$ with $e(H) \geq e(G)/r$ contains a P_{n+1}.*

Note that any improvement of the order of magnitude of p in the above theorem would improve the upper bound for $\hat{R}(P_n, r)$. However, Theorem 4(ii) implies that when $p \ll (\log cr)/(6n)$, i.e. $(\log cr)/np \gg 6$, a.a.s. there exists a P_{n+1}-free subgraph of $G \in G(crn, p)$ which contains more than

$$\frac{1}{40}\frac{(\log cr)/np}{\log((\log cr)/np)}pn \cdot crn \geq cpn \cdot crn > e(G)/r$$

edges. Therefore, $(\log r)/n$ is the threshold function for the density statement in Theorem 12. It would be interesting to know if $(\log r)/n$ is still the threshold function for the corresponding Ramsey-type statement.

5. Concluding remarks

Our investigation raises some open problems. The most interesting question is to investigate the corresponding Ramsey properties on random graphs. The Ramsey-type questions on sparse random graphs have been studied by several researchers, for example, see [7,31].

Problem 13. Determine the threshold function $p(n)$ for the following statement. For some constant c and $r \geq 2$ (c is independent of r), every r-coloring of $G(crn, p)$ contains a monochromatic P_{n+1}.

Theorem 12 implies that $p(n) = O((\log r)/n)$, while the lower bound of $\hat{R}(P_n, r)$ shows that $p(n) = \Omega(1/n)$, where n goes to infinity. The exact behavior of $p(n)$ remains open and its determination would be very useful for studying the size-Ramsey number of paths.

Another direction is to consider the following graph parameter. Denote by $c(G, F)$ the minimum number of colors k such that there exists a k-coloring of G without monochromatic F. Clearly, we have

$$c(G(N,p), P_{n+1}) \geq \frac{\binom{N}{2}p}{\mathrm{ex}(G(N,p), P_{n+1})} \geq \frac{pN^2}{3\mathrm{ex}(G(N,p), P_{n+1})}. \tag{7}$$

Let $r = N/n$. We first present two general upper bounds on $c(G(N,p), P_{n+1})$.

Theorem 14. *Suppose r is a prime power, then $c(G(N,p), P_{n+1}) \leq r+1$.*

Proof. We use a construction from [17] (also appeared in [26]). Let A_r be an affine plane of order r, i.e. r^2 points with $r^2 + r$ lines, where every pair of points is contained in a unique line, and the lines can be split into $r+1$ disjoint families $F_1, \ldots, F_{r+1}$ so that the lines inside the families are parallel.

We arbitrarily partition $[N]$ into r^2 parts $V_1, V_2, \ldots, V_{r^2}$, where each part has size $N/r^2 = n/r$. We define an $r+1$-coloring as follows. If e is an edge crossing between V_x and V_y, where the unique line containing xy is in the family F_i, then we color e by i. Observe that every connected subgraph in color i has its vertex set V inside $\cup_{x\in L} V_x$ for some line $L \in A_r$. Therefore, we have $|V| \leq r \cdot n/r = n$, and there is no monochromatic P_{n+1}. □

Theorem 15. *A.a.s. $c(G(N,p), P_{n+1}) \leq 2pN$.*

Proof. Let $k = 2pN$, and we can assume $k \leq r+1$. Consider a random k-coloring of $G(N,p)$. Then the subgraph G_i, whose edges are all edges in color i, is in $G(N, p')$, where $p' = p/k = 1/2N$. A fundamental result of Erdős and Rényi shows that a.a.s the largest component of G_i has size $O(\log N) \leq n$. Therefore, a.a.s. there is no monochromatic P_{n+1}. □

Corollary 16. *If $p = \frac{1}{\omega \cdot n}$, where $\omega = \omega(r) \geq 2$, then a.a.s. $c(G(N,p), P_{n+1}) \leq 2r/\omega$.*

For the lower bound, the proof of Theorem 1.2. in [12] implies the following.

Theorem 17. *For $p \geq 22(\log(r/7))/n$, a.a.s. $c(G(N,p), P_{n+1}) > r/7$.*

This together with Theorem 14 shows that a.a.s. $c(G(N,p), P_{n+1}) = \Theta(r)$ for $p = \Omega((\log r)/n)$. On the other hand, Theorem 4 and (7) give a lower bound for small p.

Theorem 18. *For $p \leq (\log r)/34n$, a.a.s. $c(G(N,p), P_{n+1}) \geq \frac{\log \omega}{24\omega} r$, where $\omega = (\log r)/np$.*

This naturally raises the following question.

Problem 19. What is the exact behavior of $c(G(N,p), P_{n+1})$ for $p = o((\log r)/n)$, where n goes to infinity?

Acknowledgments

We would like to thank to the referees for their valuable comments and suggestions.

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, The longest path in a random graph, Combinatorica 1 (1) (1981) 1–12.
[2] L. Babai, M. Simonovits, J. Spencer, Extremal subgraphs of random graphs, J. Graph Theory 14 (5) (1990) 599–622.
[3] D. Bal, L. DeBiasio, New lower bounds on the size-Ramsey number of a path, 2019, arXiv preprint arXiv:1909.06354.
[4] J. Balogh, R. Morris, W. Samotij, Independent sets in hypergraphs, J. Amer. Math. Soc. 28 (3) (2015) 669–709.
[5] J. Beck, On size Ramsey number of paths, trees, and circuits. I, J. Graph Theory 7 (1) (1983) 115–129.
[6] I. Ben-Eliezer, M. Krivelevich, B. Sudakov, The size Ramsey number of a directed path, J. Combin. Theory Ser. B 102 (3) (2012) 743–755.
[7] T. Bohman, A. Frieze, M. Krivelevich, P.-S. Loh, B. Sudakov, Ramsey games with giants, Random Struct. Algorithms 38 (1–2) (2011) 1–32.
[8] B. Bollobás, The evolution of sparse graphs, in: Graph Theory and Combinatorics, Proceedings Cambridge Combinatorial Conference in Honour of Paul Erdős, 1984, pp. 335–357.
[9] D. Conlon, W.T. Gowers, Combinatorial theorems in sparse random sets, Ann. of Math. (2) 184 (2) (2016) 367–454.
[10] D. Dellamonica Jr., Y. Kohayakawa, M. Marciniszyn, A. Steger, On the resilience of long cycles in random graphs, Electron. J. Combin. 15 (1) (2008) Research Paper 32, 26 pages.
[11] A. Dudek, P. Prałat, On some multicolor Ramsey properties of random graphs, SIAM J. Discrete Math. 31 (3) (2017) 2079–2092.
[12] A. Dudek, P. Prałat, Note on the multicolour size-Ramsey number for paths, Electron. J. Combin. 25 (3) (2018) Research Paper 3.35, 5 pages.
[13] P. Erdős, R.J. Faudree, C.C. Rousseau, R.H. Schelp, The size Ramsey number, Period. Math. Hungar. 9 (1–2) (1978) 145–161.
[14] P. Erdős, T. Gallai, On maximal paths and circuits of graphs, Acta Math. Hungar. 10 (3–4) (1959) 337–356.
[15] P. Frankl, V. Rödl, Large triangle-free subgraphs in graphs without K_4, Graphs Combin. 2 (1) (1986) 135–144.
[16] A.M. Frieze, On large matchings and cycles in sparse random graphs, Discrete Math. 59 (3) (1986) 243–256.
[17] A. Gyárfás, Partition coverings and blocking sets in hypergraphs, Commun. Comput. Autom. Inst. Hungar. Acad. Sci. 71 (1977) 62 (in Hungarian).
[18] P.E. Haxell, Y. Kohayakawa, T. Łuczak, Turán's extremal problem in random graphs: Forbidding even cycles, J. Combin. Theory Ser. B 64 (2) (1995) 273–287.
[19] P.E. Haxell, Y. Kohayakawa, T. Łuczak, Turán's extremal problem in random graphs: Forbidding odd cycles, Combinatorica 16 (1) (1996) 107–122.
[20] A. Johansson, J. Kahn, V. Vu, Factors in random graphs, Random Struct. Algorithms 33 (1) (2008) 1–28.
[21] Y. Kohayakawa, B. Kreuter, A. Steger, An extremal problem for random graphs and the number of graphs with large even-girth, Combinatorica 18 (1) (1998) 101–120.
[22] Y. Kohayakawa, T. Łuczak, V. Rödl, On K_4-free subgraphs of random graphs, Combinatorica 17 (2) (1997) 173–213.
[23] Y. Kohayakawa, V. Rödl, M. Schacht, The Turán theorem for random graphs, Combin. Probab. Comput. 13 (1) (2004) 61–91.
[24] J. Komlós, E. Szemerédi, Limit distribution for the existence of Hamiltonian cycles in a random graph, Discrete Math. 43 (1) (1983) 55–63.
[25] M. Krivelevich, Expanders—how to find them, and what to find in them, in: Surveys in Combinatorics 2019, in: London Math. Soc. Lecture Note Ser., vol. 456, Cambridge Univ. Press, Cambridge, 2019, pp. 115–142.
[26] M. Krivelevich, Long cycles in locally expanding graphs, with applications, Combinatorica 39 (1) (2019) 135–151.
[27] M. Krivelevich, G. Kronenberg, A. Mond, Turán-type problems for long cycles in random and pseudo-random graphs, 2019, arXiv preprint arXiv:1911.08539.
[28] O. Pikhurko, A note on the Turán function of even cycles, Proc. Amer. Math. Soc. 140 (11) (2012) 3687–3692.
[29] D. Saxton, A. Thomason, Hypergraph containers, Invent. Math. 201 (3) (2015) 925–992.
[30] M. Schacht, Extremal results for random discrete structures, Ann. of Math. (2016) 333–365.
[31] R. Spöhel, A. Steger, H. Thomas, Coloring the edges of a random graph without a monochromatic giant component, Electron. J. Combin. 17 (1) (2010) Research Paper 133, 7 pages.
[32] T. Szabó, V.H. Vu, Turán's theorem in sparse random graphs, Random Struct. Algorithms 23 (3) (2003) 225–234.