Combinatorics, Probability and Computing (2021), **30**, pp. 619–630
doi:10.1017/S0963548320000607

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Large complete minors in random subgraphs

Joshua Erde[1,†,*], Mihyun Kang[1,†] and Michael Krivelevich[2,‡]

[1]Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria and [2]School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
*Corresponding author. Email: erde@tugraz.at

(Received 15 April 2020; revised 5 October 2020; accepted 15 September 2020; first published online 3 December 2020)

Abstract

Let G be a graph of minimum degree at least k and let G_p be the random subgraph of G obtained by keeping each edge independently with probability p. We are interested in the size of the largest complete minor that G_p contains when $p = (1+\varepsilon)/k$ with $\varepsilon > 0$. We show that with high probability G_p contains a complete minor of order $\tilde{\Omega}(\sqrt{k})$, where the $\sim$ hides a polylogarithmic factor. Furthermore, in the case where the order of G is also bounded above by a constant multiple of k, we show that this polylogarithmic term can be removed, giving a tight bound.

2020 MSC Codes: Primary: 05C80; Secondary: 05C83

1. Introduction

The binomial random graph model $G(n, p)$, introduced by Gilbert [6], is a random variable on the subgraphs of the complete graph K_n whose distribution is given by including each edge in the subgraph independently with probability p. Since its introduction this model has been extensively studied. A particularly striking feature of this model is the 'phase transition' that it undergoes at $p = 1/n$, exhibiting vastly different behaviour when $p = (1-\varepsilon)/n$ to when $p = (1+\varepsilon)/n$ (where ε is a positive constant). For more background on the theory of random graphs, see [1], [4] and [9].

More recently, the following generalization of the binomial random graph model has attracted attention. Suppose G is an arbitrary graph with minimum degree $\delta(G)$ at least $k-1$, and let G_p denote the random subgraph of G obtained by retaining each edge of G independently with probability p. When $G = K_k$, the complete graph on k vertices, we recover the binomial model $G(k, p)$.

For several properties, it has been shown that once one passes the threshold for the occurrence of the property which holds in $G(k, p)$ with high probability[a] (as a function of k), or w.h.p. for short, these properties will also occur w.h.p. in G_p. For example, when $p = (1+\varepsilon)/k$ it has been shown that w.h.p. G_p is non-planar [5], and contains a path or cycle of length linear in k [2, 17]. Similarly, when $p = \omega(1/k)$, w.h.p. G_p contains a path or cycle of length $(1-o(1))k$ [15, 19] and

[†]Supported by the Austrian Science Fund (FWF): I3747.
[‡]Supported in part by USA–Israel BSF grant 2018267, and by ISF grant 1261/17.
[a]Here and throughout the paper, we will say that an event happens with high probability (w.h.p.) if the probability tends to one as $k \to \infty$. All asymptotics in the paper are taken as $k \to \infty$.

when $(1+\epsilon)(\log k)/k$, w.h.p. G_p contains a path of length k [15] and in fact even a cycle of length $k+1$ [7]. All of these results generalize known results about the binomial model.

In this paper we will be interested in the size of the largest complete minor in a graph G, sometimes known as the *Hadwiger number* of G, which we denote by $h(G)$. Fountoulakis, Kühn and Osthus [3] showed the following bound for the Hadwiger number of $G(k,p)$ in the so-called *supercritical regime*.

Theorem 1.1 ([3]). *Let ε be a positive constant and $p=(1+\varepsilon)/k$. Then w.h.p. $h(G(k,p))=\Theta(\sqrt{k})$.*

Using expanders, Krivelevich [13] gave an alternative proof of the above theorem.

As part of their work on the genus of random subgraphs, Frieze and Krivelevich [5] noted that their proof actually shows that if G is a graph with minimum degree at least k and $p=(1+\varepsilon)/k$, then w.h.p. $h(G_p)=\omega(1)$, and asked what the largest function $t(k)$ is such that w.h.p. $h(G_p)\geqslant t(k)$.

Our main result is a lower bound on $h(G_p)$, which is tight up to polylogarithmic factors.

Theorem 1.2. *Let ε be a positive constant, let G be a graph with $\delta(G)\geqslant k$, and $p=(1+\varepsilon)/k$. Then w.h.p.*

$$h(G_p)=\Omega\left(\sqrt{\frac{k}{\log k}}\right).$$

In other words, for any $\varepsilon>0$, there exists a constant $c=c(\varepsilon)$ and a function $f:\mathbb{N}\to[0,1]$ such that if $k\in\mathbb{N}$ is large enough, $(G^i:\ i\in\mathbb{N})$ is a sequence of graphs with $\delta(G^i)\geqslant k$, and $p=(1+\varepsilon)/k$, then

$$\mathbb{P}\left(h(G^i_p)\leqslant c\sqrt{\frac{k}{\log k}}\right)\leqslant f(k),$$

and $f(k)\to 0$ as $k\to\infty$.

Using ideas similar to the proof of Krivelevich in [13], we are able to remove the polylogarithmic factor, and to give the following asymptotically tight bound, when the number of vertices in G is linear in k.

Theorem 1.3. *Let ν and ε be positive constants, let G be a graph on n vertices with $\delta(G)\geqslant k\geqslant \nu n$, and $p=(1+\varepsilon)/k$. Then w.h.p.*

$$h(G_p)=\Omega\big(\sqrt{k}\big).$$

Note that if $k=\Theta(n)$, then w.h.p. the number of edges in G_p is at most $(1+\epsilon)n^2/k=O(n)$. Hence, since any graph with a K_t minor must contain at least $e(K_t)=\binom{t}{2}$ edges, it follows that w.h.p. $h(G_p)=O(\sqrt{n})=O(\sqrt{k})$, and so this bound is indeed asymptotically tight. We would be interested to know if this is the correct bound for all ranges of k.

Question 1.4. Let ε be a positive constant, let G be a graph with $\delta(G)\geqslant k$, and $p=(1+\varepsilon)/k$. Is $h(G_p)=\Omega(\sqrt{k})$ w.h.p.?

A key ingredient in our proof will be the following lemma, which roughly says that if we have a forest T of order n whose components are all of size around $\sqrt{k}$ and a set F of $\Theta(kn)$ edges on the same vertex set as T, and if $p=\Theta(1/k)$, then w.h.p. the random subgraph $T\cup F_p$ will contain a complete minor of order around $\sqrt{k}$.

Lemma 1.1. *Let $k = \omega(1)$ and $n = \omega(\sqrt{k})$ be integers, and let $b_1, c_1, c_2 > 0$ and $b_2 > 1$ be constants. Suppose V is a set of n vertices, T is a spanning forest of V with components $A_1, \ldots, A_r \subseteq V$ such that $b_1\sqrt{k} \leqslant |A_i| \leqslant b_2\sqrt{k}$, F is a set of c_1kn edges on the vertex set V, and $p = c_2/k$. Then w.h.p.*

$$h(T \cup F_p) = \Omega\left(\sqrt{\frac{k}{\log k}}\right).$$

The paper is structured as follows. In Section 2 we will introduce the relevant background material and some useful lemmas. In Section 3 we will give a proof of Lemma 1.1 and then in Sections 4 and 5 we will give proofs of Theorems 1.2 and 1.3.

Notation

Throughout the paper we will omit floor and ceiling signs to simplify the presentation. We will write log for the natural logarithm, and given a graph G we let $|G|$ denote the number of vertices in G.

2. Preliminaries

We will use the following bound, originally from Kostochka [11, 12] and Thomason [20], which says a graph of large average degree contains a large complete minor.

Lemma 2.1 ([21]). *If the average degree of G is at least $t\sqrt{\log t}$, then $h(G) \geqslant t$.*

Corollary 2.2. *If the average degree of G is at least t, then $h(G) = \Omega(t/\sqrt{\log t})$.*

We will also want to use the following simple lemma, which essentially appears in [16], to decompose a tree into roughly equal-sized parts.

Lemma 2.3 ([16, Proposition 4.5]). *Let T be a rooted tree on n vertices with maximum degree Δ, and let $1 \leqslant \ell \leqslant n$ be an integer. Then there exists a vertex $v \in V(T)$ such that the subtree T_v of T rooted at v satisfies $\ell \leqslant |T_v| \leqslant \ell\Delta$.*

As a corollary we have the following decomposition result for a tree with bounded maximum degree.

Corollary 2.4. *If T is a tree with $\Delta(T) \leqslant C$ and $|T| > \sqrt{k}$, then there exist disjoint vertex sets $A_1, \ldots, A_r \subseteq V(T)$ such that*

- $V(T) = \bigcup_{i=1}^{r} A_i$,
- *$T[A_i]$ is connected for each i, and*
- *$\sqrt{k} \leqslant |A_i| \leqslant (C+1)\sqrt{k}$ for each i.*

We will need the following simple bound on the expectation of a restricted binomial random variable.

Lemma 2.5. *Let $X \sim \mathrm{Bin}(n, p)$ be a binomial random variable with $2enp < K$ for some constant $K > 0$. If $Y = \min\{X, K\}$, then*

$$\mathbb{E}(Y) \geqslant np - K2^{-K}.$$

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

Proof. For every $t \leqslant K$ we have that $\mathbb{P}(Y=t) \geqslant \mathbb{P}(X=t)$. Hence, by standard estimates,

$$\begin{aligned}\mathbb{E}(X)-\mathbb{E}(Y) &\leqslant \sum_{t>K} t\binom{n}{t}p^t(1-p)^{n-t}\\ &\leqslant \sum_{t>K} t\left(\frac{enp}{t}\right)^t\\ &\leqslant \sum_{t>K} enp\left(\frac{enp}{t}\right)^{t-1}\\ &\leqslant \sum_{t>K} \frac{K}{2}\left(\frac{enp}{K}\right)^{t-1}\\ &\leqslant \frac{K}{2}\left(\frac{enp}{K}\right)^{K-1}\\ &\leqslant K2^{-K},\end{aligned}$$

since $enp/K < 1/2$. □

We will use the following generalized Chernoff-type bound, due to Hoeffding.

Lemma 2.6 ([8]). *Let $K>0$ be a constant and let $X_1,\dots,X_n$ be independent random variables such that $0 \leqslant X_i \leqslant K$ for each $i \leqslant n$. If $X=\sum_{i=1}^n X_i$ and $t \geqslant 0$, then*

$$\mathbb{P}(|X-\mathbb{E}(X)| \geqslant t) \leqslant 2\exp\left(-\frac{t^2}{nK^2}\right).$$

3. Large complete minors: proof of Lemma 1.1

Since $V=\bigcup_{i=1}^r A_i$ and $b_1k^{1/2} \leqslant |A_i| \leqslant b_2k^{1/2}$, it follows that $r \leqslant b_1^{-1}k^{-\frac{1}{2}}n$. Let F' be the set of edges in F which are not contained in any A_i. Then, since each A_i contains at most $\binom{|A_i|}{2} \leqslant b_2^2k/2$ edges inside it and $|F| \geqslant c_1kn$, it follows that for large k,

$$|F'| \geqslant |F| - r\frac{b_2^2}{2}k \geqslant c_1kn - \frac{b_2^2}{2b_1}\sqrt{k}n \geqslant \frac{c_1}{2}kn.$$

Hence on average each A_i meets at least $2|F'|/r \geqslant c_1b_1k^{3/2}$ edges in F'. We recursively delete sets A_i, and the edges in F' incident to them, which meet at most $c_1b_1k^{\frac{3}{2}}/4$ edges remaining in F'; we must eventually stop this process before exhausting the A_i, since $r \leqslant b_1^{-1}k^{-\frac{1}{2}}n$ (*i.e.* there are at most $b_1^{-1}k^{-\frac{1}{2}}n$ many A_i) and

$$\frac{c_1b_1}{4}k^{3/2}\frac{1}{b_1}k^{-1/2}n = \frac{c_1}{4}kn \leqslant \frac{|F'|}{2}.$$

Hence there is some subfamily, without loss of generality, $\{A_1,\dots,A_\ell\}$ of the A_i, and some subset $F'' \subseteq F'$ of edges which lie between A_i and A_j with $i,j \in [\ell]$ such that at least $c_1b_1k^{\frac{3}{2}}/4$ edges of F'' meet each A_i.

Note that $0 \leqslant e_{F''}(A_i,A_j) \leqslant b_2^2k$ for each pair $i,j \in [\ell]$. For each pair $i,j \in [\ell]$ such that $e_{F''}(A_i,A_j) > k$, let us delete $e_{F''}(A_i,A_j)-k$ edges in F'' which lie between A_i and A_j, and call the resulting set of edges $\hat{F}$. Then $0 \leqslant e_{\hat{F}}(A_i,A_j) \leqslant k$ for each $i,j \in [\ell]$, and furthermore each A_i still meets at least $c_1b_1b_2^{-2}k^{\frac{3}{2}}/4$ edges of $\hat{F}$. Indeed, the proportion of the edges in F'' between each pair A_i and A_j that we delete is at most $1-b_2^{-2}$ proportion of the edges meeting each A_i remains. In particular we have

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

$$\sum_{i,j\in[\ell]} e_{\hat{F}}(A_i, A_j) \geqslant \ell \frac{c_1 b_1}{2b_2^2} k^{3/2}. \tag{3.1}$$

Let H be an auxiliary (random) graph on $[\ell]$ such that $i \sim j$ if and only if there is an edge between A_i and A_j in $\hat{F}_p$. The number of edges between A_i and A_j in $\hat{F}_p$ is distributed as $\mathrm{Bin}(e_{\hat{F}}(A_i, A_j), p)$. Note that if $mp < 1/2$, then

$$\mathbb{P}(\mathrm{Bin}(m, p) \neq 0) = 1 - (1-p)^m \geqslant \frac{mp}{2}.$$

Since $e_{\hat{F}}(A_i, A_j) \leqslant k$ and $p = c_2/k$, and without loss of generality we may assume that $c_2 < 1/2$, it follows that

$$\mathbb{P}(i \sim j) \geqslant \frac{c_2 e_{\hat{F}}(A_i, A_j)}{2k}. \tag{3.2}$$

By (3.1) and (3.2), we have

$$\mathbb{E}(e(H)) = \frac{1}{2} \sum_{i,j\in[\ell]} \mathbb{P}(i \sim j) \geqslant \frac{1}{2} \sum_{i,j\in[\ell]} \frac{c_2 e_{\hat{F}}(A_i, A_j)}{2k} \geqslant \frac{1}{4k} \ell \frac{c_1 c_2 b_1}{2b_2^2} k^{3/2} = \frac{c_1 c_2 b_1}{8b_2^2} \ell k^{1/2}.$$

Summing up, we have $v(H) = \ell$ and $\mathbb{E}(e(H)) = \Omega(\ell k^{1/2})$, and so we expect H to have average degree $\Omega(k^{1/2})$. It remains to show that $e(H)$ is well concentrated about its mean $\mu := \mathbb{E}(e(H))$.

Since $e(H)$ can be expressed as the sum of independent indicator random variables, a standard calculation shows that $\mathrm{Var}(e(H)) \leqslant \mu$ and so, by Chebyshev's inequality,

$$\mathbb{P}(|e(H) - \mu| \geqslant \mu^{2/3}) \leqslant \frac{\mathrm{Var}(e(H))}{\mu^{4/3}} \leqslant \mu^{-1/3} = o(1).$$

Hence w.h.p. $e(H) \geqslant (1 - o(1))\mu$ and so w.h.p. H has average degree $\Omega(k^{1/2})$. Thus, by Corollary 2.2, w.h.p.

$$h(H) = \Omega\left(\sqrt{\frac{k}{\log k}}\right).$$

Observe that by contracting each A_i the graph H becomes a minor of $T \cup F_p$, and so the result follows.

4. The general case: Proof of Theorem 1.2

We will broadly follow the strategy of Frieze and Krivelevich [5] and their proof that w.h.p. G_p is non-planar when $\delta(G) \geqslant k$ and $p = (1+\varepsilon)/k$. Using a lemma similar to Lemma 1.1, they showed that if there is a tree T in G_{p_1}, where

$$p_1 = \frac{1 + \varepsilon/2}{k},$$

with small maximum degree and $\Omega(|T|k)$ edges in G, then, after exposing these edges with probability $p_2 \geqslant \varepsilon/(2k)$, the resulting graph will w.h.p. be non-planar. Since by Corollary 2.4 we can split such a tree into components of size around $\sqrt{k}$, we can use Lemma 1.1 in a similar fashion to find a large complete minor in this case.

In order to find such a tree, Frieze and Krivelevich first build a small tree T_1 with small maximum degree, and then in stages iteratively expose the edges leaving the frontier S_t (*i.e.* the set of active leaves) of the current tree T_t under the assumption that $|S_t| = \Theta(|T_t|)$ and that the maximum degree in T_t is small (in their argument polylogarithmic in k).

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

If many of the edges leaving S_t go back into the tree T_t, then we can apply Lemma 1.1 as above to find a large complete minor. Otherwise, many of the edges leave T_t, in which case Frieze and Krivelevich showed that one can either find a dense subgraph between S_t and its neighbourhood, and so also a large complete minor by Theorem 1.3, or add a new layer of significant size to the current tree, whilst keeping the maximum degree bounded, allowing one to grow a slightly larger tree. Since this process cannot continue indefinitely, as G is finite, eventually the tree stops growing and we find our large minor.

However, one cannot guarantee that the dense subgraph one finds is particularly dense, and so following this strategy naively only produces a minor of size *logarithmic* in k. Instead, by exposing (the edges emanating from) the vertices of S_t *sequentially*, we will show that if we cannot continue the tree growth, then at some point during the process there are many edges in G between the new layer of growth and the remaining vertices in S_t, allowing us to apply Lemma 1.1 as before.

Proof of Therorem 1.2. Our plan will be to sprinkle with

$$p_1 = \frac{1+\varepsilon/2}{k} \quad \text{and} \quad p_2 = \frac{p-p_1}{1-p_1} \geqslant \frac{\varepsilon}{2k}.$$

Initial phase

We first run an initial phase in which we build a partial binary tree T_0 of size $\log\log\log k =: N$ or $N+1$ in G_{p_1}. By a partial binary tree we mean a rooted tree, rooted at a leaf ρ, in which all vertices have degree three or one, such that there is some integer L such that every non-root leaf is at distance L or $L-1$ from ρ.

We will do so via a sequence of trials. In a general stage we will have a set of *discarded vertices* X which will have size $o(\log k)$, and a partial binary tree T' of size $< N$, such that so far we have only exposed edges in G_{p_1} which meet either X, the root of T', or a non-leaf vertex of T'.

If T' is a single vertex, let v be the root of T'; otherwise let $v \in V(T')$ be a non-root leaf of minimal distance to the root. We expose the edges between v and $V \backslash (X \cup V(T')$ in G_{p_1}. If v has at least two neighbours, we choose two of them arbitrarily and add them to T' as children of v, choosing and adding only one if v is the root of T'. Otherwise we say that the trial *fails* and we add $V(T')$ to X and choose a new root v arbitrarily from $V \backslash X$ and set $T' = v$. If at any point $|T'| = N$ or $N+1$, we set $T_0 := T'$ and we finish the initial phase.

Since each v has at least $k - |X \cup V(T')| \geqslant (1-\varepsilon)k$ neighbours in $V \backslash (X \cup V(T'))$, the probability that a trial fails is at most

$$\begin{aligned}
\mathbb{P}(\mathrm{Bin}((1-\varepsilon)k, p_1) < 2) &= (1-p_1)^{(1-\varepsilon)k} + (1-\varepsilon)kp_1(1-p_1)^{(1-\varepsilon)k-1} \\
&\leqslant \left(1 - p_1 + (1-\varepsilon)\left(1+\frac{\varepsilon}{2}\right)\right) \exp\left(-\left(1+\frac{\varepsilon}{2}\right)(1-\varepsilon) + p_1\right) \\
&\leqslant 2e^{\varepsilon - 1} =: 1 - \gamma < 1.
\end{aligned}$$

Since each successful trial, apart from the first, adds two new vertices to T', each time we choose a new root the probability that we build a suitable T_0 before a trial fails is at least γ^N.

Therefore w.h.p. we build such a tree before we have chosen $\gamma^{-N}N$ new roots. Since we only ever discard at most N vertices, during this process the number of discarded vertices is at most

$$\gamma^{-N}N^2 = (\log\log k)^{-\log\gamma}(\log\log\log k)^2 = o(\log k).$$

Let S_0 be the set of non-root leaves of T_0. Since T_0 is a partial binary tree as defined above, T_0 is contained in a full binary tree of depth L rooted at ρ, and so $|T_0| \leqslant 2^L$, and since all of its non-root leaves are at depth $L-1$ or L, it follows that $|S_0| \geqslant 2^{L-2}$. In particular, $|S_0| \geqslant |T_0|/4$. Furthermore, during this process we have only exposed edges which are incident to either a vertex in X or a vertex in $V(T_0) \backslash S_0$. In particular, we have not exposed any edges between S_0 and $V \backslash (X \cup V(T_0))$.

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

Tree branching phase

Suppose then that in a general step we have a tree T_t together with a set S_t of leaves of T_t, called the *frontier* of T_t, with the following properties:

(1) $|S_t| \geqslant \epsilon|T_t|/16$,
(2) no edges from S_t to $V\backslash(X \cup V(T_t))$ have been exposed in G_{p_1},
(3) the maximum degree in T_t is at most $K+1$, where

$$K := 4\log\frac{1}{\varepsilon}$$

is a large constant.

Note that T_0 and S_0 satisfy these three properties.

Let $0 < \delta \ll \varepsilon$ and let us consider the set

$$V_0 = V_0(t) := \{s \in S_t \colon e_G(s, T_t) \geqslant \delta k\}.$$

If $|V_0| \geqslant \delta|S_t|$, then $G[V(T_t)]$ contains a set F of at least

$$\frac{\delta^2}{2}|S_t|k \geqslant \frac{\delta^2\varepsilon}{32}|T_t|k$$

edges. In particular, note that this implies that $|T_t| = \Omega(k)$.

Since T_t has bounded degree, by Corollary 2.4 we can split it into connected pieces of size $\Theta(\sqrt{k})$, and hence by Lemma 1.1, when we sprinkle onto the edges of F with probability p_2, w.h.p. we obtain a complete minor of order $\Omega(\sqrt{k/\log k})$.

So we may assume that $|V_0| \leqslant \delta|S_t|$. Let $V_1 = V_1(t) := S_t\backslash V_0$. Since $|X| = o(k)$, every vertex $s \in V_1$ has degree at least $(1-2\delta)k$ to $V\backslash(X \cup V(T_t))$. Let us arbitrarily order the set $V_1 = \{s_1, \ldots, s_r\}$ where $r := |V_1|$.

We will build the new frontier S_{t+1} by exposing the neighbourhood of each s_i in turn. At the start of the process each s_i has at least $(1-2\delta)k$ possible neighbours; however, as S_{t+1} grows, it may be that some s_i have a significant fraction of their neighbours inside S_{t+1}.

Let us initially set $S_{t+1}(0) = \emptyset$ and $B(0) = \emptyset$. We will show that w.h.p. we can either find a large complete minor, or construct, for each $1 \leqslant j \leqslant r$, sets $S_{t+1}(j)$ and $B(j)$, and a forest $F(j)$, such that:

(1) $B(j) \subseteq \{s_i \colon i \in [j]\}$ and $|B(j)| < \delta|S_t|$,
(2) each $s \in B(j)$ has $e_G(s, S_{t+1}(j)) \geqslant \delta k$,
(3) there is a forest $F(j)$ of maximum degree K in G_{p_1}, whose components are stars centred at vertices in $\{s_i \colon i \in [j]\}$, such that $F(j)$ contains every vertex of $S_{t+1}(j)$.

Clearly this is satisfied with $j=0$. Suppose we have constructed appropriate $S_{t+1}(j-1)$ and $B(j-1)$.

If $d_G(s_j, S_{t+1}(j-1)) \geqslant \delta k$, then we let $B(j) = B(j-1) \cup s_j$, $S_{t+1}(j) = S_{t+1}(j-1)$ and $F(j) = F(j-1)$. If $|B(j)| \geqslant \delta|S_t|$, then we can apply Lemma 1.1 to the edges spanned by $V(T_t \cup F(j))$; those include the edges in $E_G(B(j), S_{t+1}(j))$.

By our assumptions $T_t \cup F(j)$ has bounded maximum degree, and so by Corollary 2.4 we can split it into connected parts of size around $\sqrt{k}$. Furthermore, $|T_t \cup F(j)| \leqslant |T_t| + K|S_t| = \Theta(|T_t|)$ and

$$\left|E\big(G[V(T_t \cup F(j))]\big)\right| \geqslant e_G(B(j), S_{t+1}(j)) \geqslant \delta^2|S_t|k = \Theta(|T_t|k).$$

Hence, by Lemma 1.1, after sprinkling onto $G[V(T_t \cup F(j))]$, with probability p_2 w.h.p. we have a complete minor of order $\Omega(\sqrt{k/\log k})$.

Therefore we may assume that $|B(j)| < \delta|S_t|$, and so conditions (1)–(3) are satisfied by $B(j)$, $S_{t+1}(j)$ and $F(j)$.

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

So we may assume that $d_G(s_j, S_{t+1}(j-1)) \leqslant \delta k$, and hence s_j has at least $(1-3\delta)k$ neighbours in $V \setminus (V(T_t) \cup S_{t+1}(j-1))$. We expose the neighbourhood $N(j)$ of s_j in $V \setminus (V(T_t) \cup S_{t+1}(j-1))$ in G_{p_1}. Let us choose an arbitrary subset $N'(j) \subseteq N(j)$ of size $\min\{N(j), K\}$ and let $F'(j)$ be the set of edges from s_j to $N'(j)$. We set $B(j) = B(j-1)$, $S_{t+1}(j) = S_{t+1}(j-1) \cup N'(j)$ and $F(j) = F(j-1) \cup F'(j)$. It is clear that these now satisfy (1)–(3).

Hence we may assume that we have constructed $S_{t+1}(r)$, $B(r)$ and $F(r)$. Let us set $S_{t+1} = S_{t+1}(r)$ and $T_{t+1} = T_t \cup F(r)$. Note that S_{t+1} is the frontier of T_{t+1}, so property (2) is satisfied. Furthermore, since $F(r)$ has maximum degree K, property (3) is satisfied.

Finally, we note that since $|B(r)| < \delta |S_t|$, we exposed the neighbourhood $N(j)$ of at least $(1-2\delta)|S_t|$ of the vertices in S_t. Furthermore, the size of the union of their neighbourhoods stochastically dominates a sum of restricted binomial random variables. More precisely, if we let

$$Y \sim \min\{\mathrm{Bin}((1-3\delta)k, p_1), K\},$$

then the sizes of the neighbourhoods $(N'(i)\colon i \notin B(r))$ stochastically dominate a sequence of $r - |B(r)|$ mutually independent copies of Y, $(Y_i\colon i \notin B(r))$. Hence, if we let $Z = \sum_{i \notin B(r)} Y_i$, then $|S_{t+1}|$ stochastically dominates Z.

Note that

$$1 + \frac{\varepsilon}{3} \leqslant (1-3\delta)kp_1 = (1-3\delta)\left(1 + \frac{\varepsilon}{2}\right) \leqslant 2.$$

Hence, since

$$K = 4 \log \frac{1}{\varepsilon} \geqslant 2e(1-3\delta)kp_1,$$

Lemma 2.5 implies that

$$\begin{aligned}
\mathbb{E}(Y) &\geqslant \left(1 + \frac{\varepsilon}{3}\right) - K2^{-K} \\
&\geqslant \left(1 + \frac{\varepsilon}{3}\right) - Ke^{-K/2} \\
&= \left(1 + \frac{\varepsilon}{3}\right) - 4\log\left(\frac{1}{\varepsilon}\right)\varepsilon^2 \\
&\geqslant 1 + \frac{\varepsilon}{4},
\end{aligned}$$

as long as ε is sufficiently small.

Since $r - |B(r)| \geqslant (1-2\delta)|S_t|$, it follows that

$$\mathbb{E}(Z) \geqslant (1-2\delta)|S_t|\mathbb{E}(Y) \geqslant \left(1 + \frac{\varepsilon}{5}\right)|S_t|,$$

and so by Lemma 2.6 we have that

$$\begin{aligned}
\mathbb{P}\left(|S_{t+1}| < \left(1 + \frac{\varepsilon}{8}\right)|S_t|\right) &\leqslant \mathbb{P}\left(Z < \left(1 + \frac{\varepsilon}{8}\right)|S_t|\right) \\
&\leqslant \mathbb{P}\left(|Z - \mathbb{E}(Z)| > \frac{\varepsilon}{20}|S_t|\right) \\
&\leqslant 2\exp\left(-\frac{\varepsilon^2 |S_t|^2}{400(r - |B(r)|)K^2}\right) \\
&= e^{-\Omega(|S_t|)}, \qquad (4.1)
\end{aligned}$$

since $r \leqslant |S_t|$. It follows that with probability at least $1 - e^{-\Omega(|S_t|)}$,

$$|S_{t+1}| \geqslant \left(1 + \frac{\varepsilon}{8}\right)|S_t|,$$

and it is then a simple check that

$$|S_{t+1}| \geqslant \frac{\varepsilon}{16}|T_{t+1}|$$

and hence property (1) is also satisfied.

Hence we have shown that in the tth step we can either find a large complete minor, or with probability at least $1 - e^{-\Omega(|S_t|)}$ we can continue our tree growth. However, since G is finite the tree growth cannot continue forever, and so, unless the tree growth fails at some step, we must eventually find a large minor.

Recall that the probability of failure is $o(1)$ in the initial phase, and by (4.1) the probability that the tree growth fails at some step is at most

$$\sum_t e^{-\Omega(|S_t|)} = o(1),$$

since

$$|S_0| \geqslant \frac{1}{4}\log\log\log k \quad \text{and} \quad |S_t| \geqslant \left(1 + \frac{\varepsilon}{8}\right)|S_{t-1}|.$$

Hence the total probability of failure is $o(1)$, and so w.h.p. G_p contains a large minor. □

5. The dense case: proof of Theorem 1.3

We will need some auxiliary concepts and results to prove Theorem 1.3.

Definition 5.1. Let $\alpha > 0$ be given. A graph G on n vertices is an *α-expander* if, for every set of vertices $U \subseteq V(G)$ with $|U| \leqslant n/2$, the external neighbourhood of U, denoted by $N_G(U)$, satisfies

$$|N_G(U)| \geqslant \alpha|U|.$$

The following is given as a corollary of Theorem 8.4 in [14].

Lemma 5.1 ([14]). *If G is an α-expander on n vertices with bounded maximum degree, then $h(G) = \Omega(\sqrt{n})$.*

We note that it follows from results announced in [10] that the conclusion holds without the bounded maximum degree assumption.

Definition 5.2. Let $c_1 > c_2 > 1$ and let $\beta > 0$. A graph G is *(c_1, c_2, β)-locally sparse* if

- $e(G) \geqslant c_1|G|$, and
- for every $U \subseteq V(G)$ such that $|U| \leqslant \beta|G|$, we have $e_G(U) \leqslant c_2|U|$.

Lemma 5.2 ([13, Theorem 1.1]). *Let G be a (c_1, c_2, β)-locally sparse graph on n vertices with maximum degree Δ. Then G contains an induced subgraph on βn vertices which is a γ-expander for some positive $\gamma = \gamma(c_1, c_2, \beta, \Delta)$.*

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

Proof of Therorem 1.3. Let

$$p_1 = \frac{1+\varepsilon/2}{k} \quad \text{and} \quad p_2 = \frac{p-p_1}{1-p_1} \geqslant \frac{\varepsilon}{2k}.$$

We will first give a series of claims about typical properties of G_{p_1}, which together with Lemmas 5.1 and 5.2 will imply the theorem, and then give proofs of the claims.

Firstly, we claim that there exists a constant $c_1 > 0$ such that w.h.p. there is some component C_0 of G_{p_1} with at least $c_1 k$ vertices.

Claim 1 ([18, Theorem 4]). *With high probability G_{p_1} contains a connected component C_0 with at least $\epsilon^2 k/5$ vertices.*

Next we claim that w.h.p. every large component in G_{p_1} spans many edges in G.

Claim 2. *There exists a constant $c_2 = c_2(c_1, \varepsilon, \nu) > 0$ such that w.h.p. for every connected component C of G_{p_1} of order at least $c_1 k$ we have $e_G(C) \geqslant c_2 k|C|$.*

As a consequence of Claim 2, w.h.p. the component C_0 with at least $\epsilon^2 k/5$ vertices (from Claim 1) spans many edges in G. More precisely, we have that w.h.p. $e_G(C_0) \geqslant c_2 k|C_0|$ and so, by the Chernoff bound, w.h.p. after we sprinkle with probability $p_2 \geqslant \varepsilon/(2k)$ into C_0, we have

$$e(G_p[C_0]) \geqslant |C_0| + \frac{c_2\varepsilon|C_0|}{4} \geqslant \left(1+\frac{c_2\varepsilon}{4}\right)|C_0| =: c_3|C_0|. \tag{5.1}$$

Let $c_4 := 1 + c_2\varepsilon/8$, noting that $c_3 > c_4 > 1$.

Claim 3. *There exists a constant $\beta = \beta(c_4, \varepsilon, \nu) > 0$ such that w.h.p. for every $U \subseteq V(G)$ of size $|U| \leqslant \beta k$ we have $e_{G_p}(U) \leqslant c_4|U|$.*

It follows from (5.1) and Claim 3 that w.h.p. $G_p[C_0]$ is (c_3, c_4, β)-locally sparse.

We shall show that the effect of vertices of large degree on all these estimates is small, so we can assume that $G_p[C_0]$ has bounded maximum degree. To do this we use a result from [13], which says that w.h.p. no small set of vertices meets too many edges.

Claim 4 ([13, Proposition 2]). *If $\mu > 0$ is sufficiently small and $f(\mu) = -\mu \log \mu$, then w.h.p. every set of at most $f(\mu)n$ vertices in G_p touches at most μn edges.*

Note that $f(\mu) \to 0$ as $\mu \to 0$. Let Y be the $f(\mu)n$ vertices of highest degree in G_p. If Claim 4 holds, then all vertices in $G_p \backslash Y$ have degree at most $2\mu/f(\mu)$, since otherwise the vertices in Y would meet more than

$$\frac{1}{2} f(\mu) n \left(2\frac{\mu}{f(\mu)}\right) = \mu n$$

edges in G_p, contradicting the claim. Hence w.h.p. in $G' = G_p \backslash Y$ the vertex set $C' = C_0 \backslash Y$ will span at least

$$c_3|C_0| - \mu n = \left(1 + \frac{c_2\varepsilon}{4}\right)|C_0| - \mu n$$

edges. Since $|C_0| \geqslant c_1 k \geqslant c_1 \nu n$, if $\mu(c_1, c_2, \nu)$ is sufficiently small, then there will be at least

$$\left(1 + \frac{c_2\varepsilon}{5}\right)|C_0| \geqslant \left(1 + \frac{c_2\varepsilon}{5}\right)|C_0 \backslash Y| =: c_3'|C_0 \backslash Y|$$

edges in G'. Note that $c_3' > c_4 > 1$. Furthermore, every set of at most βk vertices in G' is also a subset of G_p and so has at most $c_4|U|$ edges.

It follows that w.h.p. G' is (c_3', c_4, β)-locally sparse and its maximum degree is bounded above by $2\mu/f(\mu)$. Hence, by Lemma 5.2, w.h.p. G' contains a linear-sized (in n) expander with bounded maximum degree, and so by Lemma 5.1, w.h.p. G' (and hence G_p) contains a complete minor of order $\Omega(\sqrt{n}) = \Omega(\sqrt{k})$. □

It remains to prove Claims 2 and 3.

Proof of Claim 2. Let us say a component C of G_{p_1} is *bad* if $|C| \geqslant c_1 k$ and $e_G(C) < c_2 k|C|$. There are at most $\sum_{r \geqslant c_1 k} \binom{n}{r}$ possible vertex sets for bad components C. Furthermore, for each such set C with $|C| = r$, we have

$$\begin{aligned}
\mathbb{P}(C \text{ is a bad component}) &\leqslant \mathbb{P}(C \text{ is a component in } G_{p_1} | e_G(C) < c_2 kr) \\
&\leqslant \binom{c_2 kr}{r-1} p_1^{r-1} \\
&\leqslant (2ec_2 k)^{r-1} \left(\frac{1+\varepsilon/2}{k} \right)^{r-1} \\
&= \left(2ec_2 \left(1 + \frac{\varepsilon}{2} \right) \right)^{r-1}.
\end{aligned}$$

Hence, by the union bound, we have

$$\begin{aligned}
\mathbb{P}(\text{there exists a bad } C) &\leqslant \sum_{r=c_1 k}^{n} \binom{n}{r} \left(2ec_2 \left(1 + \frac{\varepsilon}{2} \right) \right)^{r-1} \\
&\leqslant \sum_{r=c_1 k}^{n} \left(\frac{en}{c_1 k} \right)^r \left(2ec_2 \left(1 + \frac{\varepsilon}{2} \right) \right)^{r-1} \\
&= \sum_{r=c_1 k}^{n} \left(\frac{e}{c_1 \nu} \right)^r \left(2ec_2 \left(1 + \frac{\varepsilon}{2} \right) \right)^{r-1} = o(1),
\end{aligned}$$

as long as $c_2 = c_2(c_1, \varepsilon, \nu)$ is sufficiently small. □

Proof of Claim 3. The proof goes via the union bound as above. We say a subset U is *bad* if $|U| \leqslant \beta k$ but $e_{G_p}(U) \geqslant c_4|U|$. Then we have

$$\begin{aligned}
\mathbb{P}(\text{there exists a bad } U) &\leqslant \sum_{r=1}^{\beta k} \binom{n}{r} \binom{\binom{r}{2}}{c_4 r} p^{c_4 r} \\
&\leqslant \sum_{r=1}^{\beta k} \left(\frac{en}{r} \left(\frac{er(1+\varepsilon)}{k} \right)^{c_4} \right)^r \\
&= \sum_{r=1}^{\beta k} \left(e(e(1+\varepsilon))^{c_4} \nu^{-1} \left(\frac{r}{k} \right)^{c_4 - 1} \right)^r \\
&\leqslant \sum_{r=1}^{\beta k} (e(e(1+\varepsilon))^{c_4} \nu^{-1} \beta^{c_4 - 1})^r,
\end{aligned}$$

which will be $o(1)$ as long as $\beta = \beta(c_4, \varepsilon, \nu)$ is sufficiently small. □

https://doi.org/10.1017/S0963548320000607 Published online by Cambridge University Press

Acknowledgement

Part of this work was performed when the third author visited the Institute of Discrete Mathematics at TU Graz, and this visit was financially supported by TU Graz within the Doctoral Programme 'Discrete Mathematics'. He is grateful to the Institute of Discrete Mathematics for its hospitality. We would also like to thank the reviewers for their suggestions and comments.

References

[1] Bollobás, B. (2001) *Random Graphs*, Vol. 73 of Cambridge Studies in Advanced Mathematics, second edition. Cambridge University Press.
[2] Ehard, S. and Joos, F. (2018) Paths and cycles in random subgraphs of graphs with large minimum degree. *Electron. J. Combin.* **25** 2–31.
[3] Fountoulakis, N., Kühn, D. and Osthus, D. (2008) The order of the largest complete minor in a random graph. *Random Struct. Alg.* **33** 127–141.
[4] Frieze, A. and Karoński, M. (2016) *Introduction to Random Graphs*. Cambridge University Press.
[5] Frieze, A. and Krivelevich, M. (2013) On the non-planarity of a random subgraph. *Combin. Probab. Comput.* **22** 722–732.
[6] Gilbert, E. N. (1959) Random graphs. *Ann. Math. Statist.* **30** 1141–1144.
[7] Glebov, R., Naves, H. and Sudakov, B. (2017) The threshold probability for long cycles. *Combin. Probab. Comput.* **26** 208–247.
[8] Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**(301) 13–30.
[9] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience.
[10] Kawarabayashi, K. and Reed, B. (2010) A separator theorem in minor-closed classes. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pp. 153–162. IEEE.
[11] Kostochka, A. V. (1982) The minimum Hadwiger number for graphs with a given mean degree of vertices. *Metody Diskretnogo Analiza* **38** 37–58.
[12] Kostochka, A. V. (1984) Lower bound of the Hadwiger number of graphs by their average degree. *Combinatorica* **4** 307–316.
[13] Krivelevich, M. (2018) Finding and using expanders in locally sparse graphs. *SIAM J. Discrete Math.* **32** 611–623.
[14] Krivelevich, M. (2019) Expanders: how to find them, and what to find in them. In *Surveys in Combinatorics 2019*, Vol. 456 of London Mathematical Society Lecture Note Series, pp. 115–142. Cambridge University Press.
[15] Krivelevich, M., Lee, C. and Sudakov, B. (2015) Long paths and cycles in random subgraphs of graphs with large minimum degree. *Random Struct. Algorithms* **46** 320–345.
[16] Krivelevich, M. and Nachmias, A. (2006) Coloring complete bipartite graphs from random lists. *Random Struct. Algorithms* **29** 436–449.
[17] Krivelevich, M. and Samotij, W. (2014) Long paths and cycles in random subgraphs of $\mathcal{H}$-free graphs. *Electron. J. Combin* **21** 1–30.
[18] Krivelevich, M. and Sudakov, B. (2013) The phase transition in random graphs: a simple proof. *Random Struct. Algorithms* **43** 131–138.
[19] Riordan, O. (2014) Long cycles in random subgraphs of graphs with large minimum degree. *Random Struct. Algorithms* **45** 764–767.
[20] Thomason, A. (1984) An extremal function for contractions of graphs. *Math. Proc. Camb. Phil. Soc.* **95** 261–265.
[21] Thomason, A. (2001) The extremal function for complete minors. *J. Combin. Theory Ser. B* **81** 318–338.

Cite this article: Erde J, Kang M and Krivelevich M (2021). Large complete minors in random subgraphs. *Combinatorics, Probability and Computing* **30**, 619–630. https://doi.org/10.1017/S0963548320000607