Bull. Aust. Math. Soc. **96** (2017), 1–13
doi:10.1017/S0004972716001295

BIPARTITE SUBGRAPHS OF *H*-FREE GRAPHS

QINGHOU ZENG✉ and JIANFENG HOU

(Received 23 October 2016; accepted 1 November 2016; first published online 7 February 2017)

Abstract

For a graph G, let $f(G)$ denote the maximum number of edges in a bipartite subgraph of G. For an integer m and for a fixed graph H, let $f(m, H)$ denote the minimum possible cardinality of $f(G)$ as G ranges over all graphs on m edges that contain no copy of H. We give a general lower bound for $f(m, H)$ which extends a result of Erdős and Lovász and we study this function for any bipartite graph H with maximum degree at most $t \geq 2$ on one side.

2010 *Mathematics subject classification*: primary 05C35; secondary 05C70.

Keywords and phrases: H-free graph, partition, bipartite subgraph.

1. Introduction

All graphs considered here are finite, undirected and have no loops and no parallel edges, unless otherwise indicated. All logarithms are to the natural base e. For a graph G, let $f(G)$ be the maximum number of edges in a bipartite subgraph of G. For an integer m, let $f(m)$ denote the minimum value of $f(G)$ as G ranges over all graphs with m edges.

It is easy to see that $f(m) \geq m/2$, for instance by considering a random bipartition or a suitable greedy algorithm of a graph with m edges. Edwards [9] improved the lower bound and showed that for every m,

$$f(m) \geq \frac{m}{2} + \frac{1}{4}\left(\sqrt{2m + \frac{1}{2}} - \frac{1}{4}\right). \tag{1.1}$$

Note that this is tight when $m = \binom{n}{2}$ for odd integers n. For more information on $f(m)$, including a determination of its precise value for some values of m, we refer the reader to [1, 3, 7]. For survey articles, see [8, 17].

The situation is more complicated if we consider only H-free graphs G, that is, graphs G that contain no copy of a fixed graph H. Let $f(m, H)$ denote the minimum possible cardinality of $f(G)$ as G ranges over all H-free graphs on m edges. Alon *et al.* [2] gave the following general conjecture.

This work is supported by NSFC (Grant No. 11671087) and New Century Programming of Fujian Province (Grant No. JA14028).

CONJECTURE 1.1 (Alon *et al.* [2]). For any fixed graph H, there exists a positive constant $\epsilon = \epsilon(H)$ such that

$$f(m, H) \geq \frac{m}{2} + \Omega(m^{3/4+\epsilon}).$$

Clearly, it suffices to prove this conjecture for complete graphs H. The problem of estimating the error term more precisely is not easy, even for relatively simple graphs H. The case $H = K_3$, in which $f(m, K_3)$ is the minimum possible size of the maximum cut in a triangle-free graph with m edges, has been studied extensively. Erdős and Lovász (see [10]) proved by probabilistic methods that

$$f(m, K_3) \geq \frac{m}{2} + cm^{2/3}\left(\frac{\log m}{\log\log m}\right)^{1/3}$$

for some positive constant c. After a series of papers by various researchers [16, 19], Alon [1] proved that $f(m, K_3) = m/2 + \Theta(m^{4/5})$ for all m.

In this paper, we use the method of Poljak and Tuza [16] to extend the result of Erdős and Lovász for graphs containing no copy of the complete graph K_{k+1}, and establish the following lower bound.

THEOREM 1.2. *For any fixed integer $k \geq 2$ and all $m > 1$, there exists a positive constant $c(k)$ such that*

$$f(m, K_{k+1}) \geq \frac{m}{2} + c(k)m^{k/(2k-1)}\left(\frac{\log^2 m}{\log\log m}\right)^{(k-1)/(2k-1)}.$$

Denote by $K_{t,s}$ the complete bipartite graph with classes of vertices of sizes t and s. Alon *et al.* [5] proposed a stronger conjecture for $K_{t,s}$-free graphs.

CONJECTURE 1.3 (Alon *et al.* [5]). For all $s \geq t \geq 2$ and all m, there exists a positive constant $c(s)$ such that

$$f(m, K_{t,s}) \geq \frac{m}{2} + c(s)m^{(3t-1)/(4t-2)}.$$

If true, this is tight at least for all $s \geq (t-1)! + 1$, as shown by the projective norm graphs [6]. For the cases $t = 2, 3$, the authors established the following theorem.

THEOREM 1.4 (Alon *et al.* [5]). *For $t \in \{2, 3\}$ and $s \geq t$, there exists a positive constant $c(s)$ such that*

$$f(m, K_{t,s}) \geq \frac{m}{2} + c(s)m^{(3t-1)/(4t-2)}$$

for all m, and this result is tight up to the value of $c(s)$.

In addition, Alon *et al.* [5] studied the function $f(m, H)$ for some other special bipartite graphs H.

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

THEOREM 1.5 (Alon *et al.* [5]). *Let H denote the union of an arbitrary number of cycles of length 4, all having a single common vertex. Then there exists a positive constant $c(H)$ such that*

$$f(m, H) \geq \frac{m}{2} + c(H)m^{5/6}$$

for all m, and this result is tight up to the value of $c(H)$.

In this paper, we consider the function $f(m, H)$ for any bipartite graph H with maximum degree $t \geq 2$ on one side and prove the following results.

THEOREM 1.6. *Let $H = H[X, Y]$ be a bipartite graph with vertex degree at most $t \geq 2$ for each vertex in Y.*

(i) *For each $t \geq 2$ and all m, there exists a positive constant $c(H)$ such that*

$$f(m, H) \geq \frac{m}{2} + c(H)m^{t/(2t-1)}.$$

(ii) *For $t = 2$ and all m, there exists a positive constant $c'(H)$ such that*

$$f(m, H) \geq \frac{m}{2} + c'(H)m^{5/6}.$$

(iii) *Suppose that $d(x) = |Y|$ for some vertex $x \in X$. For $t = 3$ and all m, there exists a positive constant $c''(H)$ such that*

$$f(m, H) \geq \frac{m}{2} + c''(H)m^{4/5}.$$

REMARK 1.7. Note that (i) gives a general weak lower bound in Conjecture 1.3 by setting $H = K_{t,s}$ for all $s \geq t \geq 2$. The ideas of Poljak and Tuza [16] can be used to improve the bound in (i) by logarithmic factors by more careful calculations. Finally, Theorems 1.4 and 1.5 are corollaries of (ii) and (iii).

2. K_{k+1}-free graphs

2.1. Independence numbers. In this subsection, we aim to bound the independence number $\alpha(G)$ of a K_{k+1}-free graph G in terms of its number of vertices. We need the following lemmas.

LEMMA 2.1 (Turán; see [21]). *Let G be a graph on n vertices with average degree at most d. Then*

$$\alpha(G) \geq \frac{n}{1+d}.$$

LEMMA 2.2 (Shearer [18]). *Let G be a triangle-free graph on n vertices with average degree $d > 1$. Then*

$$\alpha(G) \geq \frac{d\log d - d + 1}{(d-1)^2}n \geq \frac{\log d - 1}{d}n.$$

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

Lemma 2.3 (Li *et al.* [15]). *Let G be a graph on n vertices with average degree at most d. If the average degree of the subgraph induced by the neighbourhood of any vertex is at most a, then*

$$\alpha(G) \geq nF_{a+1}(d),$$

where

$$F_a(x) = \int_0^1 \frac{(1-t)^{1/a}}{a+(x-a)t}\,dt > \frac{\log(x/a)-1}{x} \quad (x > 0).$$

Lemma 2.4. *Let $l(x) = \log x/x$ for $x > 0$ and $L(x) = (l(\log x))^{-1}$ for $x > e$. The function $l(x)$ is monotonically increasing for $0 < x \leq e$ and decreasing for $x > e$, and the function $g(x) = L(x)/x$ is decreasing for $x > e$.*

Having finished the necessary preparations, we establish the following theorem.

Theorem 2.5. *For any fixed integer $k \geq 2$, let G be a K_{k+1}-free graph on n vertices with average degree at most d. Then*

$$\alpha(G) \geq \frac{1}{4k^2} n^{1/k} (\log n)^{(k-1)/k}.$$

Proof. Let G be a graph with maximum degree Δ. Denote by G' the graph induced by the neighbourhood of any vertex of G with maximum degree Δ and denote by G'' the graph induced by the neighbourhood of any vertex of G' with maximum degree Δ' in G'. Note that G' is K_k-free and G'' is K_{k-1}-free for $k \geq 3$.

We prove the theorem by induction on k. Let $k = 2$. Since vertex neighbourhoods in a triangle-free graph are independent sets, we may assume that $\Delta < (n \log n)^{1/2}$. If $d \leq e^2$, by Lemma 2.1,

$$\alpha(G) \geq \frac{n}{1+e^2} \geq \frac{1}{16}(n \log n)^{1/2},$$

as required. Suppose that $e^2 < d < (n \log n)^{1/2}$. From Lemmas 2.2 and 2.4,

$$\alpha(G) \geq \frac{\log d - 1}{d} n \geq \frac{\log d}{2d} n \geq \frac{1}{16}(n \log n)^{1/2}.$$

Thus, we get the desired result and establish the base case.

Assume that the result holds for any K_r-free graph with $r \leq k$ and $k \geq 3$. We show that the desired result holds for K_{k+1}-free graphs.

Claim 2.6.

$$3k^2\left(\frac{n}{\log n}\right)^{(k-1)/k} \leq d \leq \Delta \leq (n^{k-1}\log n)^{1/k} \quad \textit{and} \quad \Delta' \leq (n^{k-2}\log^2 n)^{1/k}.$$

If $d < 3k^2(n/\log n)^{(k-1)/k}$, then, by Lemma 2.1,

$$\alpha(G) \geq \frac{n}{3k^2(n/\log n)^{(k-1)/k}+1} \geq \frac{1}{4k^2} n^{1/k} (\log n)^{(k-1)/k}.$$

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

If $\Delta > (n^{k-1}\log n)^{1/k}$, then we use the induction hypothesis on G' to deduce that

$$\alpha(G) \geq \alpha(G') \geq \frac{1}{4(k-1)^2}\Delta^{1/(k-1)}(\log\Delta)^{(k-2)/(k-1)} > \frac{1}{4k^2}n^{1/k}(\log n)^{(k-1)/k}.$$

In the same way, if $\Delta' > (n^{k-2}\log^2 n)^{1/k}$, then we can also get the required result by using the induction hypothesis on G''. This completes the proof of Claim 2.6.

CLAIM 2.7.

$$\log d - \log(\Delta'+1) - 1 \geq \frac{\log d}{k}.$$

This is trivial if $\Delta' \leq 1$. Suppose that $\Delta' \geq 2$. It follows that $\Delta'+1 \leq 3\Delta'/2$. Since $\log n \geq 3k^2$ by Claim 2.6 and $k \geq 3$,

$$\frac{(n/\log n)^{(k-1)/k}}{(n^{k-2}\log^2 n)^{(1/k)\cdot(k/(k-1))}} = \left(\frac{n}{(\log n)^{k^2+1}}\right)^{1/(k^2-k)} \geq \left(\frac{e^{3k^2}}{(3k^2)^{k^2+1}}\right)^{1/(k^2-k)} \geq \frac{(3e/2)^{k/(k-1)}}{3k^2}.$$

This together with Claim 2.6 yields

$$d \geq 3k^2\left(\frac{n}{\log n}\right)^{(k-1)/k} \geq \left(\frac{3}{2}e(n^{k-2}\log^2 n)^{1/k}\right)^{k/(k-1)} \geq \left(\frac{3}{2}e\Delta'\right)^{k/(k-1)} \geq (e(\Delta'+1))^{k/(k-1)},$$

implying the desired result. This completes the proof of Claim 2.7.

By Lemmas 2.3 and 2.4 and Claim 2.7,

$$\alpha(G) \geq nF_{\Delta'+1}(d) > \frac{\log d - \log(\Delta'+1) - 1}{d}n \geq \frac{n\log d}{kd} \geq \frac{1}{4k^2}n^{1/k}(\log n)^{(k-1)/k},$$

where the last inequality follows from the fact that $e \leq d \leq (n^{k-1}\log n)^{1/k}$ by Claim 2.6. This completes the proof of Theorem 2.5. □

2.2. Chromatic numbers. In this subsection, we give an upper bound for the chromatic number $\chi(G)$ of a K_{k+1}-free graph G in terms of its number of edges.

A graph property is called *monotone* if it holds for all subgraphs of a graph with the property, that is, it is preserved under deletion of edges and vertices. We require a general lemma on monotone properties of Jensen and Toft [13] (see also [14]).

LEMMA 2.8 (Jensen and Toft [13, Section 7.3]). *For $s \geq 1$, let $\psi : [s,\infty) \to (0,\infty)$ be a positive continuous nondecreasing function. Suppose that $\mathcal{P}$ is a monotone class of graphs such that $\alpha(G) \geq \psi(|V(G)|)$ for every $G \in \mathcal{P}$ with $|V(G)| \geq s$. Then, for every such G with $|V(G)| \geq s$,*

$$\chi(G) \leq s + \int_s^{|V(G)|} \frac{1}{\psi(x)}\,dx.$$

The following lemma is an immediate corollary of Theorem 2.5 and Lemma 2.8.

LEMMA 2.9. *For any fixed integer $k \geq 2$, let G be a K_{k+1}-free graph with n vertices. Then*

$$\chi(G) \leq 16k^2\left(\frac{n}{\log n}\right)^{(k-1)/k}.$$

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

PROOF. Note that the desired result holds trivially for $n < e^2$. Suppose that $n \geq e^2$. For $x \geq e^2$, define

$$\gamma(x) = 1 - \log^{-1} x \quad \text{and} \quad \psi(x) = \frac{1}{4k^2} x^{1/k} (\log x)^{(k-1)/k}.$$

Clearly, $\gamma(x) \geq 1/2$ for $x \geq e^2$, and $\gamma(x)$, $\psi(x)$ are positive, continuous and nondecreasing. By Theorem 2.5, $\alpha(G) \geq \psi(n)$. It follows from Lemma 2.8 that

$$\chi(G) \leq e^2 + \int_{e^2}^{n} \frac{1}{\psi(x)}\, dx \leq e^2 + \frac{4k^2}{\gamma(e^2)} \int_{e^2}^{n} \frac{\gamma(x)}{x^{1/k}(\log x)^{(k-1)/k}}\, dx \leq 16k^2 \Big(\frac{n}{\log n}\Big)^{(k-1)/k},$$

where the last inequality holds because an antiderivative for the integrand is exactly $(k/(k-1))(x \log^{-1} x)^{(k-1)/k}$. Thus, we complete the proof of Lemma 2.9. □

LEMMA 2.10 (Shearer [20]). *For any fixed integer $k \geq 2$, let G be a K_{k+1}-free graph with n vertices and average degree $d > e$. Then there exists a constant $b_k \in (0, 1/4)$ such that*

$$\alpha(G) \geq \frac{b_k n \log d}{d \log \log d}.$$

The following result plays a key role in our proof of Theorem 1.2.

THEOREM 2.11. *For any fixed integer $k \geq 2$, let G be a K_{k+1}-free graph with $m > 1$ edges. Then*

$$\chi(G) \leq 32k(k + b_k^{-1}) \Big(\frac{m \log \log m}{\log^2 m}\Big)^{(k-1)/(2k-1)}.$$

PROOF. Let G be a K_{k+1}-free graph on n vertices with $m > 1$ edges. If $\chi(G) \leq 50$, then we are done. Suppose that $\chi(G) > 50$. We may also assume that G is vertex-critical. It follows that the minimal degree of G is at least 50. Thus, we have $m \geq 25n$.

For convenience, define

$$n^* = \Big(\frac{m^k \log^k \log m}{\log m}\Big)^{1/(2k-1)}.$$

We may assume that $n > n^*$. For, otherwise, $n \leq n^*$. Since $m > 1$, we see that $n \geq 3 > e$. It follows from Lemmas 2.4 and 2.9 that

$$\chi(G) \leq 16k^2 \Big(\frac{n}{\log n}\Big)^{(k-1)/k} \leq 16k^2 \Big(\frac{n^*}{\log n^*}\Big)^{(k-1)/k} \leq 32k^2 \Big(\frac{m \log \log m}{\log^2 m}\Big)^{(k-1)/(2k-1)},$$

where the last inequality follows from the fact that $x \geq 2 \log x$ for each $x > 0$. Thus, we get the desired result.

Now, we construct a graph sequence $\mathcal{G} = \{G_i\}_{i \geq 0}$ according to the following procedure, which we call the $\mathcal{G}$ algorithm. Set $i = 0$, $G_0 = G$ and $n_0 = |V(G_0)|$. Repeat the following steps until $n_i \leq n^*$:

(a) choose S_i to be the maximum independent set of G_i;

(b) set $G_{i+1} = G_i \backslash S_i$, $n_i = |V(G_i)|$ and increment i.

Let $\ell + 1$ be the length of the resulting sequence $\mathcal{G}$. By the $\mathcal{G}$ algorithm, we immediately see that $n_\ell \le n^*$ and G can be coloured by at most $\chi(G_\ell) + \ell$ colours. Since G_ℓ is K_{k+1}-free, by Lemmas 2.4 and 2.9, for $n_\ell \ge 3$,

$$\chi(G_\ell) \le 16k^2\Big(\frac{n_\ell}{\log n_\ell}\Big)^{(k-1)/k} \le 16k^2\Big(\frac{n^*}{\log n^*}\Big)^{(k-1)/k} \le 32k^2\Big(\frac{m\log\log m}{\log^2 m}\Big)^{(k-1)/(2k-1)}.$$

Clearly, the last inequality holds for $\chi(G_\ell)$ with $n_\ell \le 2$ as well. In the following, it suffices to bound the value of ℓ.

Define $t = \lceil n/n^* \rceil$. Note that $t \ge 2$ since $n > n^*$. Let $I = \{0, 1, \ldots, \ell - 1\}$ and $J = \{2, 3, \ldots, t\}$. Note that $n_i > n^* \ge n/t$ for each $i \in I$ by the $\mathcal{G}$ algorithm and the definition of t. Thus, for each $j \in J$, we can define

$$V_j = \Big\{x \in S_i : \frac{n}{j} < n_i \le \frac{n}{j-1},\ i \in I\Big\} \quad \text{and} \quad I_j = \Big\{i \in I : \frac{n}{j} < n_i \le \frac{n}{j-1}\Big\}.$$

Claim 2.12. *For each $i \in I_j \ne \emptyset$,*

$$|S_i| \ge \frac{b_k n^2}{2j^2 m} \cdot L\Big(\frac{2jm}{n}\Big),$$

where $L(x)$ is defined as in Lemma 2.4.

Let d_i denote the average degree of G_i for each $i \in I$. Clearly, for each $i \in I_j$, we have $d_i \le 2m/n_i \le 2jm/n$. Suppose that $d_i > e$. By Lemmas 2.4 and 2.10,

$$|S_i| \ge b_k n_i \cdot \frac{L(d_i)}{d_i} \ge \frac{b_k n^2}{2j^2 m} \cdot L\Big(\frac{2jm}{n}\Big),$$

as required. Otherwise, $d_i \le e$. From Lemma 2.1, $|S_i| \ge n_i/4 \ge n/(4j)$, which, together with the fact that $x \ge L(x)$ for $x > e$ and $4b_k < 1$, implies the required result as well. This completes the proof of Claim 2.12.

For each $x \in S_i$ and $i \in I$, define $w(x) = |S_i|^{-1}$. Now, for each $x \in S_i \subset V_j$, it follows from Claim 2.12 that

$$w(x) = |S_i|^{-1} \le \frac{2j^2 m}{b_k n^2 L(2jm/n)} \le \frac{2j^2 m \log\log m}{b_k n^2 \log(2jm/n)}, \tag{2.1}$$

where the last inequality holds because $j \le t \le n/2$ by the definitions of t and n^*. By the definitions of $w(x)$ and V_j,

$$\ell = \sum_{i\in I}\sum_{x\in S_i} w(x) = \sum_{j\in J}\sum_{x\in V_j} w(x) \quad \text{and} \quad |V_j| < \frac{n}{j-1} - \frac{n}{j}. \tag{2.2}$$

In view of (2.1) and (2.2),

$$\ell \le \sum_{j=2}^{t} \frac{2j^2 |V_j| m \log\log m}{b_k n^2 \log(2jm/n)} \le \frac{4m}{b_k n} \sum_{j=2}^{t} \frac{\log\log m}{\log j + \log(m/n)}. \tag{2.3}$$

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

By the definition of n^*,

$$\frac{n}{n^*} \cdot \frac{m}{n} = \frac{m}{n^*} = \left(\frac{m^{k-1} \log m}{\log^k \log m} \right)^{1/(2k-1)}. \tag{2.4}$$

It follows that

$$\max\left\{ \log \frac{m}{n}, \log \frac{n}{n^*} \right\} > \frac{\log m}{4}. \tag{2.5}$$

Suppose that $n/n^* < m/n$. Note that $t - 1 < n/n^*$ by the definition of t. Thus, we delete the first term of the denominator of (2.3) and obtain

$$\ell \leq \frac{4m}{b_k n} \sum_{j=2}^{t} \frac{\log \log m}{\log(m/n)} \leq \frac{4m \log \log m}{b_k n^* \log(m/n)} \leq \frac{16k}{b_k} \left(\frac{m \log \log m}{\log^2 m} \right)^{(k-1)/(2k-1)},$$

where the last inequality follows from (2.4) and (2.5). Otherwise, $n/n^* \geq m/n$. Since $t - 1 < n/n^* \leq t$,

$$\sum_{j=2}^{t} \frac{1}{\log j} \leq \int_2^t \frac{1}{\log x}\, dx \leq \frac{2(t-1)}{\log t} < \frac{2n}{n^* \log(n/n^*)}.$$

Deleting the second term of the denominator in (2.3) yields

$$\ell \leq \frac{4m}{b_k n} \sum_{j=2}^{t} \frac{\log \log m}{\log j} \leq \frac{8m \log \log m}{b_k n^* \log(n/n^*)} \leq \frac{32k}{b_k} \left(\frac{m \log \log m}{\log^2 m} \right)^{(k-1)/(2k-1)}.$$

Thus, we get the desired result by noting that $\chi(G) \leq \chi(G_\ell) + \ell$. This completes the proof of Theorem 2.11. □

2.3. Bipartite subgraphs of K_{k+1}-free graphs. In this short subsection, we give a proof of Theorem 1.2. We need the following lemma.

LEMMA 2.13 [1]. *Let G be a graph with m edges and chromatic number at most χ. Then*

$$f(G) \geq \frac{\chi + 1}{2\chi} m.$$

PROOF OF THEOREM 1.2. Set $c(k) = (64k^2 + 64kb_k^{-1})^{-1}$. The result now follows immediately from Lemma 2.13 and Theorem 2.11. □

3. Graphs with forbidden bipartite subgraphs

In this section, we consider the function $f(m, H)$ when H is a bipartite graph with maximum degree $t \geq 2$ on one side. We shall use the following upper bound, proved by Alon *et al.* [4], on the maximum number of edges in an H-free graph.

LEMMA 3.1 (Alon *et al.* [4]). *Let H be a bipartite graph with maximum degree $t \geq 2$ on one side. Then there exists a positive constant $c = c(H)$ such that*

$$ex(n, H) \leq cn^{2-1/t}.$$

A graph is *r-degenerate* if every one of its subgraphs contains a vertex of degree at most r. We need the following well-known fact (see [1, 2] or [5] for a proof).

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

LEMMA 3.2. *Let H be an r-degenerate graph on h vertices. Then there is an ordering $v_1, \ldots, v_h$ of the vertices of H such that for every $1 \le i \le h$ the vertex v_i has at most r neighbours v_j with $j < i$.*

We also require the following three lemmas establishing lower bounds for $f(G)$ for graphs G in terms of different parameters.

LEMMA 3.3 (Erdős *et al.* [11]). *Let G be a graph on n vertices with m edges and positive minimum degree. Then*

$$f(G) \ge \frac{m}{2} + \frac{n}{6}.$$

LEMMA 3.4 (Alon *et al.* [5]). *There exist two small constants $\epsilon, \delta \in (0, 1)$ such that the following holds. Let G be a graph on n vertices with m edges and degree sequence $d_1, d_2, \ldots, d_n$. Suppose, further, that for each i the induced subgraph on all the d_i neighbours of vertex number i contains at most $\epsilon d_i^{3/2}$ edges. Then*

$$f(G) \ge \frac{m}{2} + \delta \sum_{i=1}^{n} \sqrt{d_i}.$$

LEMMA 3.5 (Alon [1]). *Let $G = (V, E)$ be a graph with m edges. Suppose that $U \subset V$ and let G' be the induced subgraph of G on U. If G' has m' edges, then*

$$f(G) \ge f(G') + \frac{m - m'}{2}.$$

Finally, we shall employ a martingale concentration result to prove the existence of certain induced subgraphs in a graph with relatively large minimum degree and sparse neighbourhood.

LEMMA 3.6 (Janson *et al.* [12, Corollary 2.27]). *Given positive real numbers λ, $C_1, \ldots, C_n$, let $f : \{0, 1\}^n \to \mathbb{R}$ be a function satisfying the following Lipschitz condition: whenever two vectors $z, z' \in \{0, 1\}^n$ differ only in the ith coordinate (for any i), we always have $|f(z) - f(z')| \le C_i$. Suppose that $X_1, \ldots, X_n$ are independent random variables, each taking values in $\{0, 1\}$. Then the random variable $Y = f(X_1, \ldots, X_n)$ satisfies*

$$\mathbb{P}(|Y - \mathbb{E}[Y]| \ge \lambda) \le 2 \exp\left\{ - \frac{\lambda^2}{2 \sum_{i=1}^{n} C_i} \right\}.$$

Now, we use this to control the performance of a randomised induced subgraph of a given graph with properties stated as before.

THEOREM 3.7. *Let $G = (V, E)$ be a graph on n vertices with m edges and minimum degree at least m^θ for some fixed real $\theta \in (0, 1)$. Suppose that m is sufficiently large and the induced subgraph on the neighbourhood of any vertex $v \in V$ of degree d_v contains fewer than $sd_v^{3/2}$ edges for some positive constant s. Then, for every constant $\eta \in (0, 1)$, there exists an induced subgraph $G' = (V', E')$ of G with the following properties:*

(a) *G' contains at least $\eta^2 m/2$ edges;*
(b) *every vertex v of degree d_v in G that lies in V' has degree at least $\eta d_v/2$ in G';*
(c) *every neighbourhood of the vertex v in V' contains at most $2\eta^2 sd_v^{3/2}$ edges in G'.*

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

Proof. For each vertex $v \in V$, denote by d_v the degree of v in G and denote by e_v the number of edges of H_v induced by $N_G(v)$. Write $S = \{v \in V : e_v > 2\eta^2 s d_v^{3/2}\}$.

Let $\eta \in (0, 1)$ be any fixed real number. Consider a random subset V' of V obtained by picking each vertex of V randomly and independently, with probability η. Let G' be the subgraph of G induced by V'. Define the random variables X and Y_v to be the number of edges of G' and the degree of v in G', respectively. Thus,

$$\mathbb{E}[X] = \eta^2 m \quad \text{and} \quad \mathbb{E}[Y_v] = \eta d_v.$$

Clearly, flipping the assignment of $v \in V$ cannot affect X by more than d_v, and switching the choice of a single vertex $u \in N_G(v)$ can only change Y_v by at most 1. Define

$$L = \sum_{v\in V} d_v^2 \leq 2mn \quad \text{and} \quad L_v = d_v.$$

By Lemma 3.6,

$$\mathbb{P}\Big(X \leq \mathbb{E}[X] - \frac{1}{2}\eta^2 m\Big) \leq 2\exp\Big\{-\frac{\eta^4 m^2}{8L}\Big\} \leq 2\exp\Big\{-\frac{\eta^4 m}{16n}\Big\} \tag{3.1}$$

and

$$\mathbb{P}\Big(Y_v \leq \mathbb{E}[Y_v] - \frac{1}{2}\eta d_v\Big) \leq 2\exp\Big\{-\frac{\eta^2 d_v^2}{8L_v}\Big\} \leq 2\exp\Big\{-\frac{\eta^2 d_v}{8}\Big\}. \tag{3.2}$$

Now, we define the random variable Z_v to be the number of edges induced by $N_{G'}(v)$. Clearly, we have $\mathbb{E}[Z_v] = \eta^2 e_v$. For each $v \in S$, switching the choice of a single $u \in N_G(v)$ can only affect Z_v by at most $d_{H_v}(u)$. Similarly, if we define $L'_v = \sum_{u\in N_G(v)} d^2_{H_v}(u) \leq 2e_v d_v$, then Lemma 3.6 gives

$$\mathbb{P}(Z_v \geq \mathbb{E}[Z_v] + \eta^2 e_v) \leq 2\exp\Big\{-\frac{\eta^4 e_v^2}{2L'_v}\Big\} \leq 2\exp\Big\{-\frac{\eta^6 s\sqrt{d_v}}{2}\Big\}. \tag{3.3}$$

Note that $d_v \geq m^\theta$ for each $v \in V$ and some fixed real $\theta \in (0, 1)$. Since $2m = \sum_{v\in V} d_v$, we have $m = \Omega(n)$. Thus, each of (3.1)–(3.3) holds with probability exponentially small in n for sufficiently large m. Since there are at most $2n + 1$ conditions to check and each fails with probability exponentially small in n, some choice of V' has the required properties. This completes the proof of Theorem 3.7. □

Proof of Theorem 1.6. (i) Let H be a bipartite graph with maximum degree $t \geq 2$ on one side and let $G = (V, E)$ be an H-free graph with n vertices and m edges. By Lemma 3.1, there exists a constant $c_1 = c_1(H) > 1$ such that $m \leq c_1 n^{2-1/t}$.

Let $d(v)$ denote the degree of v in G. Define $S = \{v \in V : d(v) \geq 4c_1 n^{1-1/t}\}$. Clearly, $|S| \leq n/2$. Let G' be the subgraph of G induced by $V \backslash S$. Note that G' contains at least $n/2$ vertices and has maximum degree at most $4c_1 n^{1-1/t}$. By Lemma 2.1,

$$\alpha(G) \geq \alpha(G') \geq \frac{n/2}{1 + 4c_1 n^{1-1/t}} \geq \psi(n),$$

where $\psi(x) = (10c_1)^{-1}x^{1/t}$. Note that $\psi(x)$ is positive, continuous and nondecreasing. From Lemma 2.8,

$$\chi(G) \le 1 + \int_1^n \frac{10c_1}{x^{1/t}}\,dx \le 20c_1 n^{1-1/t}. \tag{3.4}$$

If $n \ge c_1' m^{t/(2t-1)}$ for some constant $c_1' > 1$, then Lemma 3.3 gives

$$f(G) \ge \frac{m}{2} + \frac{n}{6} \ge \frac{m}{2} + \frac{c_1'}{6} m^{t/(2t-1)}.$$

Otherwise, $n < c_1' m^{t/(2t-1)}$. In view of Lemma 2.13 and (3.4), we conclude that

$$f(G) \ge \frac{m}{2} + \frac{m}{40c_1 n^{1-1/t}} > \frac{m}{2} + \frac{1}{40c_1c_1'} m^{t/(2t-1)}.$$

Since G is chosen arbitrarily, we get the desired result by setting $c(H) = (40c_1c_1')^{-1}$. This completes the proof of (i).

(ii) Let H be a bipartite graph with maximum degree 2 on one side and let G be an H-free graph with n vertices and m edges. On account of the inequality (1.1), we may assume that m is sufficiently large. In addition, the desired result follows immediately from Lemma 3.3 for $n \ge (1/2)m^{5/6}$. Thus, we may assume that $n < (1/2)m^{5/6}$.

Claim 3.8. *There exists an induced subgraph G' of G such that G' contains at least $\eta^2 m/4$ edges and every neighbourhood of a vertex of degree d in G' spans at most $\epsilon d^{3/2}$ edges in G', where $\eta \in (0, 1)$ is a fixed constant and ϵ is a constant defined as in Lemma 3.4.*

As long as there is a vertex of degree smaller than $m^{1/6}$ in G, omit it. This process terminates after deleting fewer than $m^{1/6}n < m/2$ edges, and thus we obtain an induced subgraph $\widetilde{G}$ of G with at least $m/2$ edges and minimum degree at least $m^{1/6}$. Note that the induced subgraph on the neighbourhood of any vertex of degree $\widetilde{d}$ of $\widetilde{G}$ contains no copy of H, and hence contains at most $c_2\widetilde{d}^{\,3/2}$ edges for some constant $c_2 > 1$, by Lemma 3.1. Now, we apply Theorem 3.7 to $\widetilde{G}$ with $\eta = \epsilon^2/(32c_2^2)$. Thus, we find an induced subgraph G' of $\widetilde{G}$ (and hence of G) with the required properties. This completes the proof of Claim 3.8.

Claim 3.9. *G' is ℓ-degenerate, where $\ell = \lceil \mu m^{1/3} \rceil$ and $\mu = \mu(H) > 1$ is a fixed constant.*

Otherwise, we may assume that G' contains a subgraph G'' with minimum degree more than ℓ. Note that the number of vertices of G'' is $N < 2m/\ell \le 2\ell^2/\mu^3$. Thus, the number of edges of G'' is

$$e(G'') \ge \tfrac{1}{2}\ell N \ge (\tfrac{1}{2}\mu N)^{3/2}. \tag{3.5}$$

Since G'' is H-free and H has maximum degree 2 on one side, by Lemma 3.1, there exists a constant $c_2' = c_2'(H) > 1$ such that $e(G'') \le c_2' N^{3/2}$, which contradicts (3.5) for a chosen value $\mu > 2(c_2')^{2/3}$. This completes the proof Claim 3.9.

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press

By Lemma 3.2 and Claim 3.9, there is a labelling $v_1, v_2, \ldots, v_{n'}$ of the n' vertices of G' such that $d_i^+ \le \ell$ for every i, where d_i^+ denotes the number of neighbours v_j of v_i with $j < i$ in G'. Note that $\sum_{i=1}^{n'} d_i^+ = |E(G')|$. Let d_i be the degree of v_i in G' for each $1 \le i \le n'$. By Lemma 3.4 and Claim 3.8,

$$f(G') \ge \frac{|E(G')|}{2} + \delta \sum_{i=1}^{n'} \sqrt{d_i} \ge \frac{|E(G')|}{2} + \delta \sum_{i=1}^{n'} \sqrt{d_i^+}$$
$$\ge \frac{|E(G')|}{2} + \frac{\delta \sum_{i=1}^{n'} d_i^+}{\sqrt{\ell}} \ge \frac{|E(G')|}{2} + \frac{\delta\eta^2}{8\sqrt{\mu}} m^{5/6},$$

where $\delta = \delta(G')$ is a constant. This together with Lemma 3.5 gives

$$f(G) \ge f(G') + \frac{m - |E(G')|}{2} \ge \frac{m}{2} + \frac{\delta\eta^2}{8\sqrt{\mu}} m^{5/6}.$$

Since G is chosen arbitrarily, we get the desired result and complete the proof of (ii).

(iii) Let $H = H[X, Y]$ be a bipartite graph with vertex degree at most 3 for each vertex in Y and let G be an H-free graph with n vertices and m edges. Suppose that $d(x) = |Y|$ for some vertex $x \in X$ and denote by H' the subgraph of H induced by $(X \backslash \{x\}) \cup Y$.

With an argument similar to the one stated in the proof of (ii), we may assume that m is sufficiently large and $n < (1/2)m^{4/5}$. Note that H' is a bipartite graph with vertex degree at most 2 for each vertex in Y and the induced subgraph on the neighbourhood of any vertex in G contains no copy of H'. A similar argument to that of Claims 3.8 and 3.9, the details of which we omit, suggests the following claim.

Claim 3.10. *Let $\eta \in (0, 1)$ be a fixed constant and let ϵ be a constant defined as in Lemma 3.4. Set $\ell = \lceil \mu m^{2/5} \rceil$, where $\mu = \mu(H) > 1$ is a fixed constant. Then there is an induced subgraph G' of G with the following properties:*

(a) *G' is an ℓ-degenerate graph with at least $\eta^2 m/4$ edges;*
(b) *every neighbourhood of a vertex of degree d in G' spans at most $\epsilon d^{3/2}$ edges in G'.*

The remainder of the argument is analogous to that in (ii). By Lemmas 3.2 and 3.4 and Claim 3.10, $f(G')$ exceeds half the number of edges of G' by at least $\Omega(m^{4/5})$ and the desired result follows from Lemma 3.5. This completes the proof of (iii). □

Acknowledgement

The authors would like to thank the referees for their careful reading and helpful comments.

References

[1] N. Alon, 'Bipartite subgraphs', *Combinatorica* **16** (1996), 301–311.

[2] N. Alon, B. Bollobás, M. Krivelevich and B. Sudakov, 'Maximum cuts and judicious partitions in graphs without short cycles', *J. Combin. Theory Ser.* B **88** (2003), 329–346.

[3] N. Alon and E. Halperin, 'Bipartite subgraphs of integer weighted graphs', *Discrete Math.* **181** (1998), 19–29.

[4] N. Alon, M. Krivelevich and B. Sudakov, 'Turán numbers of bipartite graphs and related Ramsey-type questions', *Combin. Probab. Comput.* **12** (2003), 477–494.

[5] N. Alon, M. Krivelevich and B. Sudakov, 'Maxcut in H-free graphs', *Combin. Probab. Comput.* **14** (2005), 629–647.

[6] N. Alon, L. Rónyai and T. Szabó, 'Norm-graphs: variations and applications', *J. Combin. Theory Ser.* B **76** (1999), 280–290.

[7] B. Bollobás and A. D. Scott, 'Better bounds for Max Cut', in: *Contemporary Combinatorics*, Bolyai Society Mathematical Studies, 10 (Springer, Berlin, 2002), 185–246.

[8] B. Bollobás and A. D. Scott, 'Problems and results on judicious partitions', *Random Structures Algorithms* **21** (2002), 414–430.

[9] C. S. Edwards, 'An improved lower bound for the number of edges in a largest bipartite subgraph', in: *Recent Advances in Graph Theory: Proc. 2nd Czechoslovak Sympos. Graph Theory* (Academia, Praha, 1975), 167–181.

[10] P. Erdős, 'Problems and results in graph theory and combinatorial analysis', in: *Graph Theory and Related Topics: Proc. Conf. Waterloo, 1977* (Academic Press, New York, 1979), 153–163.

[11] P. Erdős, A. Gyárfás and Y. Kohayakawa, 'The size of the largest bipartite subgraphs', *Discrete Math.* **177** (1997), 267–271.

[12] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs* (John Wiley, New York, 2000).

[13] T. R. Jensen and B. Toft, *Graph Coloring Problems*, Wiley-Interscience Series in Discrete Mathematics and Optimization (John Wiley, New York, 1995).

[14] A. Kostochka, B. Sudakov and J. Verstraëte, 'Cycles in triangle-free graphs of large chromatic number', *Combinatorica* doi:10.1007/s00493-015-3262-0.

[15] Y. Li, C. Rousseau and W. Zang, 'Asymptotic upper bounds for Ramsey functions', *Graphs Combin.* **17** (2001), 123–128.

[16] S. Poljak and Zs. Tuza, 'Bipartite subgraphs of triangle-free graphs', *SIAM J. Discrete Math.* **7** (1994), 307–313.

[17] A. D. Scott, 'Judicious partitions and related problems', in: *Surveys in Combinatorics*, London Mathematical Society Lecture Note Series, 327 (Cambridge University Press, Cambridge, 2005), 95–117.

[18] J. Shearer, 'A note on independence number of triangle-free graphs', *Discrete Math.* **46** (1983), 83–87.

[19] J. Shearer, 'A note on bipartite subgraphs of triangle-free graphs', *Random Structures Algorithms* **3** (1992), 223–226.

[20] J. Shearer, 'On the independence number of sparse graphs', *Random Structures Algorithms* **7** (1995), 269–271.

[21] V. K. Wei, 'A lower bound on the stability number of a simple graph', Bell Laboratories Technical Memorandum 81-11217-9, 1981.

QINGHOU ZENG, Center for Discrete Mathematics,
Fuzhou University, Fujian, 350003, China
e-mail: qinghouzeng@hotmail.com

JIANFENG HOU, Center for Discrete Mathematics,
Fuzhou University, Fujian, 350003, China
e-mail: jfhou@fzu.edu.cn

https://doi.org/10.1017/S0004972716001295 Published online by Cambridge University Press