ELSEVIER

Discrete Mathematics 248 (2002) 125–141

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

# Disjoint subgraphs of large maximum degree

Oleg Pikhurko *,[1], Andrew Thomason

*DPMMS, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK*

Received 25 April 2000; accepted 5 February 2001

**Abstract**

Erdős [in: Chartrand (Ed.), The Theory and Applications of Graphs, Wiley, New York, 1981, p. 331] conjectured that the vertices of any graph with fewer than $\binom{2n+1}{2} - \binom{n}{2}$ edges can be split into two parts, both parts inducing subgraphs of maximum degree less than $n$. Recently, the first named author [Combinatorica 21 (2001) 403–412] disproved this conjecture. In this paper we consider further questions arising out of the conjecture.

First of all, we give couterexamples to the conjecture having only $2n+80$ vertices for large $n$. (The above counterexample had around $n^{3/2}/\sqrt{2}$ vertices, though it had many fewer edges than our examples.)

We also define the function $b(n,m)$ to be the minimum size of a graph $G$ such that, for any partition $V(G) = A \cup B$, either $\Delta(G[A]) \geqslant n$ or $\Delta(G[B]) \geqslant m$ holds. In this terminology, Erdős's conjecture was $b(n,n) = \binom{2n+1}{2} - \binom{n}{2}$. We prove that $b(n,m) = 2nm - m^2 + \mathrm{O}(\sqrt{m})n$ for $n \geqslant m, b(n,1) = 4n-2$ for $n \geqslant 7$, and $b(n,2) = 6n + \mathrm{O}(1)$.

Let $m(n,k,j)$ be the minimum size of a graph $G$ on $n+k$ vertices in which $\Delta(G[A]) \geqslant n$ for every $(n+j)$-set $A \subset V(G)$. We prove that, if $k = \mathrm{o}(n(n+j)/\log n)$, then

$$m(n,k,j) = (1+\mathrm{o}(1))\left(1 + \frac{k-j}{2n+2j}\right)(k-j+1)n$$

as $n \to \infty$. The upper bound here disproves a conjecture made by Erdős, Reid, Schelp and Staton [Discrete Math. 158 (1996) 283–286]. 

*Keywords:* Size Ramsey numbers; Bipartition; Maximum degree

## 1. Introduction

Whilst investigating size Ramsey numbers, Erdős [2] conjectured that any graph of size less than $\binom{2n+1}{2} - \binom{n}{2}$ is an edge–disjoint union of a bipartite graph and a graph

* Corresponding author.

*E-mail address:* o.pikhurko@dpmms.cam.ac.uk (O. Pikhurko).

[1] Supported by a Senior Rouse Ball Studentship, Trinity College, Cambridge, UK.


PII: S0012-365X(01)00190-X

of maximum degree less than $n$. This value arises from the consideration of the graph $P_{n+1,n}=K_{n+1}+\bar{K}_n$, which does not admit such a representation.

This conjecture has recently been disproved by the first author [6]. In fact, there are graphs with $n^2+\Theta(n^{3/2})$ edges that are not a union of the appropriate kind, and this size is the smallest possible. But the graph described in [6] has $(1/\sqrt{2}+o(1))n^{3/2}$ vertices, and one might ask what happens if not so many vertices are allowed. Counterexamples with only $3n+1$ vertices can be obtained from the methods in [6]. Clearly, all counterexamples have at least $2n+1$ vertices, and Faudree showed that they must have at least $2n+2$ (see e.g. [3] for a proof); in fact, Erdős and Faudree [4] believed that there are no counterexamples on $2n+2$ vertices. In Section 2 we show that there are counterexamples with only $2n+80$ vertices if $n$ is large.

In investigating the conjecture of Erdős, it is natural to consider also the corresponding 'off-diagonal' problem. To be precise, let $b(n,m)$ be the minimum number of edges in a graph $G$ such that, for any partition $V(G)=A\cup B$, at least one of $\Delta(G[A])\geqslant n$ and $\Delta(G[B])\geqslant m$ holds. Clearly, $b(n,n)$ is the function investigated by Erdős. In Section 3 we prove that $b(n,m)=2nm-m^2+O(\sqrt{m})n$ when $n\geqslant m$. Note that, when $n=m$, this disproves the conjecture of Erdős. For fixed values of $m$ we cannot give an asymptotic formula for $b(n,m)$ unless $m\leqslant 2$, though we offer both upper and lower bounds.

A different, but related, 'one-sided' problem was discussed by Erdős, Reid, Schelp and Staton [3]. Let $n$, $j$ and $l$ be integers with $j\geqslant 1$ and $l\geqslant 0$. The problem is to compute $q(n,j,l)$, the minimum size of a graph $G$ having $n+j+l$ vertices, such that for every $(n+j)$-set $A\subset V(G)$ we have $\Delta(G[A])\geqslant n$. (This function was denoted $m(n,k,j)$ in [3], where $k=j+l$; we changed the notation because all our arguments involve $l$ rather than $k$.) The functions $b()$ and $q()$ are closely connected: on the one hand, if $l=n-1$ then the addition of a new vertex joined to everything in $G$ shows that $b(n,n)\leqslant q(n,j,n-1)+2n-1+j$ for all $j$, and on the other hand, graphs that are extremal for the function $b(n,n)$ show that $b(n,n)\geqslant q(n,j,n)$ for some $j$.

It was proved in [3] that

$$q(n,j,l)=(l+1)n+\binom{l+1}{2} \quad \text{if } n\geqslant \max\left(jl,\binom{l+2}{2}\right). \tag{1}$$

The same equation was shown to hold whenever $j=1$, and it was conjectured [3, Conjecture 1] to hold whenever $n\geqslant j+l$. The stated value arises from the consideration of $P_{l+1,n}$ together with $j-1$ extra isolated vertices.

However, the conjecture is not true in general. We prove in Section 4 that

$$q(n,j,l)=(1+o(1))(l+1)n\left(1+\frac{l}{2n+2j}\right) \tag{2}$$

holds if $l=o(n(n+j)/\log n)$. This disproves the conjecture of Erdős, Reid, Schelp and Staton if both $j$ and $l$ are $\Theta(n)$.

The upper bound in (2) comes from a probabilistic argument. We present also a construction showing that the equation in (1) fails to hold when $n<(j-1)l$.

On the other hand, we show that the equation in (1) is true if

$$n \geqslant \left(j+\frac{1}{2}\right) l + \frac{2j+l}{4j-2},$$

which is an improvement on (1) if $j$ is smaller than about $l/2$. This shows that the value $jl$ is roughly the threshold on $n$ when the obvious construction suggesting (1) fails to be extremal.

Of course, one can also consider $q(n, j, l)$ when $l$ is much bigger than $n$ or when $n$ is fixed altogether. However, it seems that for these different ranges we encounter different kinds of problems, and the overall situation is not clear. For example, to compute $q(1, j, l)$ we must look at graphs of order $1+j+l$ with no independent set of order $1+j$; thus $q(1, j, l) = \binom{1+j+l}{2} - t_j(1+j+l)$, and the (unique) extremal graph is the complement of the corresponding Turán graph, that is, it consists of $j$ disjoint cliques of almost equal size.

## 2. Small counterexamples

The purpose of this section is to show that there are counterexamples to the conjecture of Erdős [2] having only $2n+80$ vertices, if $n$ is large. Our simple construction is based on two sparse graphs, whose existence is established in the next two lemmas by elementary probabilistic methods. The lemmas are stated in terms of explicit constants rather than in a general form, since we need only the stated versions.

**Lemma 1.** *For all large even $n$, there exists a 19-regular graph $G$ of order $n$, in which any subgraph of minimum degree 10 has more than $n/4$ vertices.*

**Proof.** The proof is an exercise in the use of the standard model for $k$-regular random graphs. A graph is generated by taking a random *configuration*, that is, a random 1-factor on $kn$ vertices, these vertices being grouped into $n$ groups of $k$ apiece. Each group is then identified to a single vertex to produce a $k$-regular multigraph. This process produces a $k$-regular graph with probability bounded away from zero (Bollobás [1, Chapter II.4]). Thus, it suffices for us to prove that a random 19-regular multigraph almost certainly satisfies the property in the lemma.

In fact, we shall do slightly more; we shall show that in almost all random 19-regular multigraphs, every subset of $l$ vertices spans fewer than $5l$ edges for all $l \leqslant n/4$. Note that $l$ vertices span at least $5l$ edges precisely when there are at most $9l$ edges between those $l$ vertices and the remaining $n-l$ vertices. Let $A$ be some collection of $l \leqslant \lfloor n/4 \rfloor$ groups, and let $B$ denote $n-l$ groups not in $A$. The total number of configurations is $\Phi(kn) = (kn)!2^{-kn/2}/(kn/2)!$, where $k=19$. The probability that a configuration has exactly $m$ edges between the $kl$ vertices in $A$ and the $k(n-l)$ vertices in $B$ (note that this implies that $kl$ and $m$ have the same parity) is

$$\binom{kl}{m}\binom{k(n-l)}{m}m!\frac{\Phi(kl-m)\Phi(k(n-l)-m)}{\Phi(kn)}$$

which is equal to

$$\binom{kn/2-m}{kl/2-m/2}\binom{kn/2}{m}\binom{kn}{kl}^{-1}2^m.$$

Denote this probability by $P(l,m)$. Thus, the probability that a random $k$-regular multigraph fails to satisfy the property in the lemma is at most $\sum_{1\leqslant l\leqslant n/4}\binom{n}{l}\sum_{m\leqslant 9l}P(l,m)$. Now

$$\frac{P(l,m+2)}{P(l,m)}=\frac{(kl-m)(kn-kl-m)}{(m+1)(m+2)}\geqslant 5 \qquad \text{for } m+2\leqslant 9l\leqslant 9n/4,$$

So $\sum_{m\leqslant 9l}P(l,m)\leqslant 2P(l,9l)$. Therefore, if we write

$$E_l=2\binom{n}{l}P(l,9l)=\binom{n}{l}\binom{kn/2-9l}{5l}\binom{kn/2}{9l}\binom{kn}{kl}^{-1}2^{9l+1},$$

then the probability of failure is bounded by $\sum_{1\leqslant l\leqslant n/4}E_l$.

We now invoke the standard estimates

$$\left(\frac{a}{b}\right)^b\leqslant\binom{a}{b}\leqslant\left(\frac{ea}{b}\right)^b \quad\text{and}\quad \binom{a}{b}=\exp\{-aH(b/a)+o(a)\},$$

where $H(x)=x\log(x)+(1-x)\log(1-x)$. These (crudely) yield the bounds

$$E_l<\frac{n^l(kn/2)^{14l}(kl)^{kl}}{l^l(5l)^{14l}(kn)^{kl}}e^{25l}\leqslant\left(\frac{k^{14}l^4}{n^4}\right)^l \quad\text{and}\quad E_{\lambda n}=\exp\{-nJ(\lambda)+o(n)\},$$

where

$$J(\lambda)=-18H(\lambda)+\left(\frac{19}{2}-9\lambda\right)H\left(\frac{10\lambda}{19-18\lambda}\right)+\frac{19}{2}H\left(\frac{18\lambda}{19}\right)-9\lambda\log 2.$$

The first bound implies $E_l=O(n^{-4})$ for $1\leqslant l\leqslant nk^{-4}$. Now it is straightforward to check that $J(\lambda)$ is a concave function with $J(0)=0$ and $J(1/4)>0$, so there exists $\varepsilon>0$ such that $J(\lambda)\geqslant\varepsilon$ for $k^{-4}\leqslant\lambda\leqslant 1/4$. Therefore $\sum_{1\leqslant l\leqslant n/4}E_l=O(n^{-3})$, completing the proof of the lemma. □

**Lemma 2.** *For all large $n$, there exists a bipartite graph with vertex classes $U$ and $W$, where $|U|=n$ and $|W|=n+79$, such that every vertex of $U$ has degree* 70, *and such that every subgraph with $n+80$ vertices containing at least $\lfloor n/4\rfloor$ vertices from $U$ has a vertex lying within $U$ of degree at most* 60.

**Proof.** Consider a random bipartite graph on the vertex sets $U$ and $W$, in which each vertex of $U$ independently chooses $d=70$ neighbours in $W$. Let $A$ be a subset of $W$ with $|A|=\lfloor n/4\rfloor-1$. Let $p$ be the probability that a given vertex in $U$ has at most 9 neighbours in $A$. Then, for large enough $n$, we have

$$p=\sum_{j=0}^{9}\binom{|A|}{j}\binom{|W|-|A|}{70-j}\binom{|W|}{70}^{-1}<\frac{20}{11}\binom{70}{9}\left(\frac{1}{4}\right)^{9}\left(\frac{3}{4}\right)^{61}<\frac{1}{90}.$$

The expected number of pairs of sets $B\subset U$ and $A\subset W$, such that $|B|=\lfloor n/4\rfloor$, $|A|=\lfloor n/4\rfloor-1$, and every vertex of $B$ has at most 9 neighbours in $A$, is thus at most, using the notation of Lemma 1,

$$\binom{n}{\lfloor n/4\rfloor}\binom{n+79}{\lfloor n/4\rfloor-1}p^{\lfloor n/4\rfloor}=\mathrm{o}(1)\times\exp\{-2nH(1/4)+(1/4)n\log p\}.$$

Since $-2H(1/4)-(1/4)\log 90<0$, there must be a graph with no such pair of sets. Any such a graph has the properties described in the lemma. □

We are now able to give a counterexample to the conjecture of Erdős.

**Theorem 3.** *For all large $n$ there is a graph $G$ of order $2n+80$ and size less than $\binom{2n+1}{2}-\binom{n}{2}-\frac{n}{2}+89$ that is not an edge–disjoint union of a bipartite graph and a graph of maximum degree less than $n$.*

**Proof.** Let $G$ have vertex set $\{v\}\cup U\cup W$, where $|U|=n$ and $|W|=n+79$. Join every vertex in $U$ to every vertex in $W$, and then delete the edges of a bipartite graph described by Lemma 2. Within $U$, join every pair of vertices and then delete the edges of a 19-regular graph described by Lemma 1 (if $n$ is odd, use a 19-regular graph of order $n+1$ with one vertex removed). Join $v$ to everything in $U\cup W$. Clearly $|G|=2n+80$, and the size of $G$ is at most

$$\binom{n}{2}-\frac{19n}{2}+\frac{19}{2}+n(n+9)+2n+79<\binom{2n+1}{2}-\binom{n}{2}-\frac{n}{2}+89$$

as desired. Consider now a partition of $V(G)$ into sets $A$ and $B$, where $v\in A$. If $|A|\geqslant n+1$ then $\Delta(G[A])\geqslant n$. If not, then $|B|\geqslant n+80$, and $B$ must contain at least one vertex of $U$. If $B$ contains at least $\lfloor n/4\rfloor$ vertices of $U$ then, by Lemma 2, one of these vertices has degree in $G[B]$ at least $|B|-1-19-60\geqslant n$. If $B$ has fewer than $\lfloor n/4\rfloor$ vertices of $U$, by Lemma 1 one of these vertices has degree in $G[B]$ at least $|B|-1-9-70\geqslant n$. Either way, $\Delta(G[B])\geqslant n$, which completes the proof. □

It is clear that the constant 80 can be reduced by analysing a little more carefully the different kinds of set $B$ that can arise in the proof of Theorem 3.

The counterexamples in Theorem 3 have only $\mathrm{O}(n)$ fewer edges than that conjectured necessary by Erdős. As explained in Section 1, there exist counterexamples with $\Omega(n^2)$ fewer edges. However, such examples must have $2n+\Omega(n)$ vertices: an estimate of how the minimum size of a counterexample varies with its order is given by Theorem 11 in Section 4.

## 3. The off-diagonal function $b(n, m)$

It is convenient to define $\mathscr{B}(n,m)$ to be the class of all graphs $G$ such that, for any partition $V(G)=A\cup B$, at least one of $\Delta(G[A])\geqslant n$ and $\Delta(G[B])\geqslant m$ holds. Thus $b(n,m)=\min\{e(G)\colon G\in\mathscr{B}(n,m)\}$. Clearly, $b(n,m)=b(m,n)$. We shall assume from now on that $n\geqslant m$. In this section we first give good bounds on $b(n,m)$ that are valid for all $n$ and $m$, and then we shall give more precise bounds for the case when $m$ is fixed and $n$ is large.

Given $v\in G$ and $A\subset V(G)$ we shall denote by $d_A(v)$ the number of neighbours of $v$ in $A$. Thus, if also $v\in A$, then $d_A(v)$ is the degree of $v$ in the induced subgraph $G[A]$.

### 3.1. General bounds

Here is a simple algorithm that gives a good general lower bound on $b(n,m)$.

Let $G\in\mathscr{B}(n,m)$. To begin with, set $A=V(G)$ and $B=\emptyset$. As long as $|B|\leqslant m$, move to $B$ any vertex $x\in A$ with $d_A(x)\geqslant n$. (Such a vertex exists, because obviously $\Delta(G[B])<m$.) When we finish, $|B|=m+1$. Now swap the sets $A$ and $B$ each with the other, so that $|A|=m+1$. If there is a vertex $y\in B$ with $d_B(y)\geqslant m$, move $y$ to $A$. Repeat this step as often as possible. Since, in the end, $\Delta(G[B])<m$, our assumption on $G$ implies that $|A|\geqslant n+1$ (to allow a vertex of degree at least $n$). Counting the edges encountered in this procedure, we obtain the following bound valid for all $n$ and $m$.

$$b(n,m)\geqslant (m+1)n+((n+1)-(m+1))m=2mn-m^2+n \tag{3}$$

Next, we provide a general construction giving an upper bound on $b(n,m)$.

Let $m=m_1+\cdots+m_f$ and $n-m=n_1+\cdots+n_g$. We define the graph $P(m_1,\ldots,m_f\,;\,n_1,\ldots,n_g)$ to be the vertex disjoint union of $P_{m_i,n}$, $1\leqslant i\leqslant f$, and $P_{n_j,m}$, $1\leqslant j\leqslant g$, together with an extra vertex $x$ joined to everything else. (Recall that $P_{s,t}=K_s+\bar{K}_t$.) Thus $P(m_1,\ldots,m_f;\ n_1,\ldots,n_g)$ has $n+1+fn+gm$ vertices, and has size

$$n+fn+gm+mn+(n-m)m+\sum_{1\leqslant i\leqslant f}\binom{m_i}{2}+\sum_{1\leqslant j\leqslant g}\binom{n_j}{2}.$$

We claim that $G=P(m_1,\ldots,m_f;\ n_1,\ldots,n_g)$ is in the class $\mathscr{B}(n,m)$. For let $V(G)$ be partitioned into two parts, $A$ and $B$. If $x\in A$, we may assume that at least $m_i$ vertices from each $P_{m_i,n}$ and at least $n_j$ vertices from each $P_{n_j,m}$ lie in $A$, for otherwise $\Delta(G[B])\geqslant m$. But then $d_A(x)=|A|-1\geqslant\sum_{1\leqslant i\leqslant f}m_i+\sum_{1\leqslant j\leqslant g}n_j=n$. On the other hand, if $x\in B$ but $\Delta(G[A])<n$, then from each $P_{m_i,n}$ at least $m_i$ vertices are in $B$, so $d_B(x)\geqslant\sum_{1\leqslant i\leqslant f}m_i=m$, as required.

Therefore $b(n,m) \leqslant e(P(m_1,\dots,m_f;n_1,\dots,n_g))$. To minimize this quantity for given $n$ and $m$, we let the $m_i$'s (and the $n_j$'s) be nearly equal while $f$ and $g$ are around $m(2n)^{-1/2}$ and $(n-m)(2m)^{-1/2}$, respectively. This, combined with the lower bound (3), yields the equation claimed in the introduction, namely the following.

**Theorem 4.** *For all* $n \geqslant m \geqslant 1$, $b(n,m) = 2nm - m^2 + O(\sqrt{m})n$ holds.

### 3.2. Fixed values of $m$

In the extreme case when $m$ is fixed and $n$ tends to infinity, consider $P(m;n_1,\dots,n_g)$ with $g = \lceil x \rceil$, where $x = n(2m)^{-1/2}$. We have

$$gm + \sum_{1 \leqslant j \leqslant g} \binom{n_j}{2} < (x+1)m + g\frac{((n-m)/x+1)n/g}{2} \leqslant n(\sqrt{2m} + 1/2),$$

so we obtain the following bounds.

**Theorem 5.** *For each fixed* $m$ *and for large* $n$, *we have*

$$(2m+1)n - m^2 \leqslant b(n,m) \leqslant (2m + \sqrt{2m} + 5/2)n - m^2/2.$$

That is, for each fixed $m \geqslant 1$, the numbers $b(n,m)$ lie between two functions linear in $n$ with slopes $2m+1$ and $2m + \sqrt{2m} + 5/2$.

The natural question to ask now is whether $\lim_{n\to\infty} b(n,m)/n$ exists and, if so, what is its value. For the cases $m=1$ and 2 we can answer this question. We prove that $b(n,1) = 4n-2$ for $n \geqslant 7$ and that $b(n,2) = 6n + O(1)$. As the reader will see, the proofs are rather lengthy. In the next case, $m=3$, the best upper bounds that we could find is $b(n,3) \leqslant 9n + O(1)$, which is obtained by considering the following example. Take the union of $\lceil n/3 \rceil$ disjoint copies of $P_{3,3}$. Choose any $n$ vertices within this union and connect them to an external copy of $K_3$. Finally, add a vertex $x$ connected to everything else. It is not hard to check that this graph has the $\mathscr{B}(n,3)$-property. However, our methods seem to give only $8n + O(1)$ as a lower bound. This indicates that making progress for other values of $m$ might be a hard task.

We begin with the case $m=1$. Write $n = 2k + l + 1$. Form the graph $G$ from the disjoint union of $k$ triangles and $l$ disjoint edges by adding two extra vertices $x$ and $y$; $x$ is connected to every other vertex (including $y$) while $y$ is connected to some $n$ vertices besides $x$, the choice being immaterial. Clearly, $e(G) = 3k + 3k + l + 2l + n + 1 = 4n - 2$.

The graph $G$ is in the class $\mathscr{B}(n,1)$. For suppose that $B$ is an independent set in $G$ and that $A = V(G) - B$. If one of $x$ or $y$ belongs to $B$, then $A$ contains the other plus their $n$ common neighbours and so $\Delta(G[A]) \geqslant n$. If $\{x,y\} \subset A$, then at least 2 vertices from each of the $k$ triangles and at least 1 vertex from each of the $l$ edges must lie in $A$ and $d_A(x) \geqslant 1 + 2k + l = n$, as required.

We shall now show that the graphs just described are extremal for $b(n,1)$. In fact, we believe that an inspection of the proof would reveal all extremal graphs to be of this kind.

**Theorem 6.** *For all* $n \geqslant 7$, $b(n,1) = 4n - 2$.

**Proof.** We have already seen that $b(n,1) \leqslant 4n-2$. So we shall assume that $G$ is a graph in the class $\mathscr{B}(n,1)$ with at most $4n-3$ edges, and we shall derive a contradiction. At various stages we shall consider sets $B \subset V(G)$; then $A$ will denote the set $V(G) - B$. Likewise if we define a subset $A \subset V(G)$ then $B$ is taken to mean $V(G) - A$. Of course, if $B$ is independent then it must be that $\Delta(G[A]) \geqslant n$, and in particular $|A| \geqslant n+1$ must hold.

Let $L$ be the set of vertices of $G$ of degree at least $n$. Certainly $L$ is not an independent set, for otherwise, taking $B = L$, we fail to have $\Delta(G[A]) \geqslant n$. In particular, $|L| \geqslant 2$. Suppose that $|L| \geqslant 4$. Any four vertices from $L$ are incident with at least $4n-6$ edges. Take $A$ to be four vertices from $L$ plus an endvertex from each edge not incident with these four vertices; there can be at most three such edges, so $|A| \leqslant 7 \leqslant n$, whilst $B$ is independent, which is a contradiction. Therefore $2 \leqslant |L| \leqslant 3$.

Suppose now that $|L| = 3$, say $L = \{x, y, z\}$. Taking $A$ to be $L$ together with an endvertex of every edge not incident with $L$, we have that $B$ is independent, so $|A| \geqslant n+1$. Thus there must be at least $n-2$ edges not incident with $L$, so at most $3n-1$ edges can meet $L$. Now if, say, $yz$ is not an edge, then taking $B = \{y, z\}$ shows that $x$ has at least $n$ neighbours outside $L$, in which case there would be at least $3n$ edges incident with $L$. So $L$ must span a triangle in $G$.

Now, taking $B$ to be any vertex of $L$, we see that there must be a different vertex in $L$ of degree at least $n+1$. Thus $L$ contains at least two vertices of degree at least $n+1$, which already implies that at least $3n-1$ edges are incident with $L$. Therefore two vertices of $L$, say $x$ and $y$, have degree exactly $n+1$, whilst $z$ has degree $n$. The neighbours of $x$ and $y$ must be identical, for if, say, $u$ is a neighbour of $x$ but not of $y$ then, taking $B = \{y, u\}$, we have $\Delta(G[A]) < n$. But now there must be a common neighbour $v$ of $x$ and $y$ that is not a neighbour of $z$, and taking $B = \{z, v\}$ we again have $\Delta(G[A]) < n$. We conclude that $|L| = 2$.

So let $L = \{x, y\}$. Since $L$ is not an independent set, $xy$ is an edge. Taking $B = \{y\}$ we see that $x$ has degree at least $n+1$, and likewise so does $y$.

We now apply a simple algorithm. To begin with, let $A$ consist of $x$ and $y$ plus all vertices connected neither to $x$ nor to $y$. We shall increase $A$ by moving vertices to it from $B$. Note that every vertex of $B$ has a neighbour in $\{x, y\}$.

At Stage 1, if there is a vertex $a \in B$ with $d_B(a) \geqslant 3$, move $a$ to $A$. Repeat as long as possible. At Stage 2, if $a \in B$ has $d_B(a) = 2$, move $a$ to $A$. Repeat as long as possible. Any edges remaining in $E(G[B])$ are now isolated. At Stage 3, if $ab \in E(G[B])$ is such that each neighbour of $a$ in $\{x, y\}$ is also a neighbour of $b$, move $a$ to $A$. Repeat as long as possible.

Suppose that $d_A(x) \geqslant n$ when this algorithm terminates. Then at least $n-1$ neighbours of $x$ were moved to $A$, and for each such neighbour $a$ we can count 3 edges of $G$ (namely, $ax$ plus two edges incident to $a$ in Stages 1 and 2, and $ax$, $ab$ and $bx$ in Stage 3). These $3(n-1)$ edges, together with $n+1$ edges incident with $y$, show that $e(G) \geqslant 4n-2$, a contradiction. Therefore $\Delta(G[A]) < n$.

Write $s_{i,j}$ for the number of vertices $a$ moved to $A$ during Stage $i$ that were incident with $j$ vertices in $\{x,y\}$, $j=1,2$; $i=1,2,3$. Also, let $t=e(G[B])$; thus $t>0$. Note that each of the $t$ edges in $E(G[B])$ induces a 4-cycle with $\{x,y\}$. One by one, for each edge $ab\in E(G[B])$ with $ax,by\in E(G)$, we could now move an endvertex to $A$ as follows: move $a$ to $A$ until $d_A(x)=n-1$, and then for the remaining edges move $b$ to $A$. Since $B$ is now independent, $y$ must now have at least $n$ neighbours in $A$. So, after the end of the original algorithm, before these $t$ vertices were moved, it must have been that

$$2n-1\leqslant t+d_A(x)+d_A(y)=t+2+\sum_{i=1}^{3}(s_{i,1}+2s_{i,2}). \tag{4}$$

There are some other edges that we have not yet taken into account. After Stage 2, $E(G[B])$ consisted of a set $I$ of $s_{3,1}+s_{3,2}+t$ isolated edges. During Stage 2, $G[B]$ had maximum degree two, and we moved out $s_{2,1}+s_{2,2}$ vertices of degree two. So there must be $r=\max\{0,s_{2,1}+s_{2,2}-s_{3,1}-s_{3,2}-t\}$ other vertices in $B$ not moved in Stages 2 or 3 nor incident with an edge of $I$. Each of these vertices has a neighbour in $\{x,y\}$. Remembering these $r$ latter edges, we clearly have the lower bound

$$e(G)\geqslant 1+r+4s_{1,1}+5s_{1,2}+3s_{2,1}+4s_{2,2}+3s_{3,1}+5s_{3,2}+3t. \tag{5}$$

Since $e(G)\leqslant 4n-3$, subtracting twice inequality (5) from inequality (4), and using the definition of $r$, we obtain

$$2\geqslant r+2s_{1,1}+s_{1,2}+s_{2,1}+s_{3,1}+s_{3,2}+t\geqslant 2s_{1,1}+s_{1,2}+2s_{2,1}+s_{2,2}.$$

Adding together these two sums, each being at most 2, then doubling the result and using (4), we find $2n-3\leqslant 8$, which contradicts $n\geqslant 7$. □

We now turn to the case $m=2$. Consider first a graph consisting of $\lceil n/2\rceil-1$ disjoint 4-cycles and one triangle, say on $X=\{x_1,x_2,x_3\}$. Form a graph $G$ from this graph by adding some further edges: join $x_1$ to every other vertex and join each of $x_2$ and $x_3$ to some set $C$ of $n$ vertices chosen from the 4-cycles, the exact choice being immaterial.

The graph $G$ is in the class $\mathscr{B}(n,2)$. To see this, take a partition of $V(G)$ into parts $A$ and $B$ with $\Delta(G[B])\leqslant 1$. If $X\cap B\neq\emptyset$, then $|(C\cup X)\cap B|\leqslant 2$, so some vertex in $X\cap A\neq\emptyset$ has at least $|C|=n$ neighbours in $A$. If $X\subset A$, then $d_A(x_1)=|A|-1\geqslant n$, because $A$ must contain at least 2 vertices from each 4-cycle. Hence, $G\in\mathscr{B}(n,2)$ and $b(n,2)\leqslant e(G)$. The graph $G$ has $6n-5$ or $6n-1$ edges according as $n\geqslant 4$ is even or odd.

We now show that the graphs just described are more or less extremal for $b(n,2)$. Given the dependence of the examples on the parity of $n$, and given also the amount of detail needed to establish the value of $b(n,1)$, we do not compute $b(n,2)$ exactly, but give instead a very sharp estimate. Even this estimate involves quite a bit of work. This tends to suggest that the complexity of the estimation of $b(n,m)$ might well increase rapidly with $m$.

**Theorem 7.** *For all* $n$, $6n-110 \leqslant b(n,2) \leqslant 6n-1$ *holds.*

**Proof.** We need only to show the lower bound. Let $G \in \mathscr{B}(n,2)$. We shall adopt the same conventions about subsets $A$ and $B$ as used in the previous proof, though in this case $\Delta(G[B]) \leqslant 1$ will imply $\Delta(G[A]) \geqslant n$.

Let $L$ be the set of vertices with degree at least $n$ in $G$. If $|L| \geqslant 4$ then let $A$ be four vertices in $L$. There are at least $4n-6$ edges currently incident with $A$. Now keep moving to $A$, as long as possible, vertices in $B$ that have degree at least two in $G[B]$. When we stop, $\Delta(G[B]) \leqslant 1$, so $|A| \geqslant n+1$ and we must have moved at least $n-3$ vertices. Therefore $e(G) \geqslant 4n-6+2(n-3) = 6n-12$, as required.

So we may assume that $|L| \leqslant 3$. But $|L| \geqslant 3$, for otherwise by taking $B = L$ we have both $\Delta(G[B]) \leqslant 1$ and $\Delta(G[A]) < n$. So $|L| = 3$; let $L = \{x, y, z\}$.

As usual, we now perform an algorithm starting with $A = L$ and $B = V(G) - L$. Moreover, we define $C \subset B$ to be the set of vertices in $G[B]$ that are neither isolated nor the endvertex of an isolated edge. During the algorithm, we keep a tally of the edges we have seen; no edge is tallied twice. We take vertices in $C$ and move them to $A$, tallying the incident edges (and sometimes others).

The idea of the proof is to carefully maintain numbers $d$ and $k$ such that $d \geqslant \max\{d_A(x), d_A(y), d_A(z)\}$ and such that the tally $t$ of edges satisfies $t \geqslant 6d - k$. Eventually we shall have $C = \emptyset$ or, equivalently, $\Delta(G[B]) \leqslant 1$, at which point we know that $d \geqslant n$, and so $e(G) \geqslant 6n - k$. Initially just the edges within $L$ are tallied, and we can take $d \leqslant 2$ and $k = 10$.

We begin by moving $\Gamma_\emptyset$ into $A$ where, for each subset $S \subset L$, $\Gamma_S$ is the set of vertices in $C$ whose neighbours in $L$ are precisely the elements of $S$. Of course, $\Gamma_S$ changes as $C$ changes, though from now on $\Gamma_\emptyset = \emptyset$ always. We shall abbreviate $\Gamma_{\{x\}}$ to $\Gamma_x$, $\Gamma_{\{x,y\}}$ to $\Gamma_{xy}$ and so on.

Whenever possible we perform the following: if $S, T \subset L$, $S \cap T = \emptyset$, and there are vertices $a \in \Gamma_S$, $b \in \Gamma_T$ with no common neighbour in $C$, then we identify $a$ and $b$ to a single vertex. This identification will not affect the number of edges incident with any other vertex in $L \cup C$, and it is immaterial if edges already tallied become identified. The identification cannot reduce $\Delta(G[B])$, nor need we increase $d$, and it remains true that $\Delta(G[B]) \leqslant 1$ will imply $d \geqslant n$.

Also whenever possible, move a vertex $a \in C$ with $d_C(a) \geqslant 5$ to $A$; at least 6 edges incident to $a$ are tallied, $d$ is increased by one and $k$ is unchanged.

Thus we now have $\Delta(G[B]) \leqslant 4$. Inside $C$, each vertex now has at most 16 vertices within distance two of it. By the remarks above, we may therefore assume that, if $S \cap T = \emptyset$ and $|\Gamma_S| > 16$, then $|\Gamma_T| = \emptyset$.

We now aim to achieve that some vertex in $L$ is joined to everything in $C$.

This is easy if, say, $|\Gamma_z| > 16$, for then $\Gamma_x = \Gamma_y = \Gamma_{xy} = \emptyset$, as desired.

We may otherwise assume that $|\Gamma_x| \leqslant |\Gamma_y| \leqslant |\Gamma_z| \leqslant 16$. Then move $\Gamma_x \cup \Gamma_y \cup \Gamma_z$ to $A$ and tally the incident edges. We increase $d$ by $|\Gamma_z|$, whilst tallying at least $3|\Gamma_z|/2$ edges, and so we increase $k$ by 72 to 82. It is now the case that $\Gamma_x = \Gamma_y = \Gamma_z = \emptyset$. Let $D = \{b \in B\colon d_B(b) \geqslant 3\}$. If $a \in \Gamma_{xyz} \cap D$, move $a$ to $A$ and increase the tally by 6, raising $d$ by one and leaving $k$ unchanged. We may repeat this until $\Gamma_{xyz} \cap D = \emptyset$. Suppose now $a \in D$, say $a \in \Gamma_{xy}$. If $\Gamma_{xz}$ has more than 4 vertices of degree at least 2 we may choose one, $b$, not adjacent to $a$. If also $\Gamma_{yz}$ has more than 8 elements, we may choose one, $c$, adjacent to neither $a$ nor $b$. Moving $\{a,b,c\}$ to $A$ and tallying their incident edges, the tally increases by at least 12, whilst $d$ increases by 2 and $k$ stays the same. Repeat this exercise until one of $\Gamma_{xy}$, $\Gamma_{xz}$ and $\Gamma_{yz}$ has at most 8 elements, or two of $\Gamma_{xy}$, $\Gamma_{xz}$ and $\Gamma_{yz}$ have at most 4 vertices of degree at least 2, or $D = \emptyset$.

If one of $\Gamma_{xy}$, $\Gamma_{xz}$ and $\Gamma_{yz}$, say $\Gamma_{xy}$, has at most 8 elements, move $\Gamma_{xy}$ to $A$. There are at least $5|\Gamma_{xy}|/2$ incident edges that we can tally, so we increase $d$ by $|\Gamma_{xy}|$ whilst increasing $k$ by 28 to 110. Since $\Gamma_x = \Gamma_y = \Gamma_{xy} = \emptyset$, everything in $C$ is joined to $z$, as desired.

If two of $\Gamma_{xy}$, $\Gamma_{xz}$ and $\Gamma_{yz}$, say, $\Gamma_{xy}$ and $\Gamma_{xz}$, each have at most 4 vertices of degree at least 2, then move these $j \leqslant 8$ vertices to $A$, tally at least $5j/2$ incident edges, increase $d$ by $j$ and increase $k$ by 28 to 110. Suppose now that $b \in \Gamma_{xy} \cup \Gamma_{xz}$. Then $d_B(b) = 1$ and (by the definition of $C$) the neighbour $a$ of $b$ in $C$ has $d_B(a) \geqslant 2$. Therefore $a \in \Gamma_{yz} \cup \Gamma_{xyz}$. Move $a$ into $A$ and tally the two edges from $b$ to $L$, the two from $a$ to $\{y,z\}$ and the edges from $a$ to $B$. Note that $b$ moves into $B - C$. The tally goes up by at least 6; increase $d$ by one and leave $k$ the same. Observe that there are still at most $d$ tallied edges incident with $x$ (though maybe more with $y$) and none of these joins $x$ to $C$. Repeat this operation until $\Gamma_{xy} \cup \Gamma_{xz} = \emptyset$; both $y$ and $z$ are now joined to everything in $C$.

So we have achieved our aim that $z$ is joined to everything in $C$, unless $\Gamma_x = \Gamma_y = \Gamma_z = \emptyset$ and $D = \emptyset$. In the latter eventuality, $\Delta(G[C]) \leqslant 2$ so $C$ comprises disjoint paths and cycles, and every vertex of $C$ has at least two neighbours in $L$. Given a path of order $p$, move $\lfloor p/3 \rfloor$ of its vertices to $A$ so that the edges remaining in $B$ are disjoint, tally $3p - 1$ edges lying in the path or joining it to $z$, and increase $d$ by $\lfloor p/3 \rfloor$. Likewise for a cycle of order $p$, move $\lceil p/3 \rceil$ vertices to $A$ so that only disjoint edges remain, and tally $3p$ edges. Since $3p - 1 \geqslant 6\lfloor p/3 \rfloor$ and $3p \geqslant 6\lceil p/3 \rceil$, $k$ need not change. But now $C = \emptyset$, so $d \geqslant n$, and we have shown $e(G) \geqslant t \geqslant 6d - k \geqslant 6n - 110$.

Finally, we may assume that $z$ is joined to everything in $C$, that $x$ is not incident with more than $d$ tallied edges, and that either the same is true for $y$ or else $y$ too is joined to everything in $C$. We write $\delta = 0$ in the first case and $\delta = 1$ in the second. There are at least $n - d$ untallied edges incident with $x$, and also with $y$ if $\delta = 0$; add them to the tally, which subsequently is at least $(2-\delta)n + (4+\delta)d - 110$. Now, if there is a vertex $a \in C$ with $d_B(a) \geqslant 3$, move it to $A$ and add to the tally the edge $az$, the edge $ay$ if $\delta = 1$, and the edges between $a$ and $B$. Increase $d$ by one, so the tally remains at least $(2-\delta)n + (4+\delta)d - 110$. Repeat this until $C$ once again consists of disjoint cycles and paths. Move vertices of $C$ to $A$, as in the previous paragraph, tallying the edges in $C$, those joining $C$ to $z$ and, if $\delta = 1$, those joining $C$ to $y$. Since $2p - 1 \geqslant 4\lfloor p/3 \rfloor$ and $2p \geqslant 4\lceil p/3 \rceil$, the tally remains at least $(2-\delta)n + (4+\delta)d - 110$. But now $d \geqslant n$, so $e(G) \geqslant 6n - 110$ as desired. □

## 4. The one-sided function $q(n, j, l)$

In this section we consider $q(n,j,l)$, the minimum size of a graph in $\mathcal{Q}(n,j,l)$, where $\mathcal{Q}(n,j,l)$ is the class of graphs $G$ having $n+j+l$ vertices, such that for every $(n+j)$-set $A \subset V(G)$ we have $\Delta(G[A]) \geqslant n$.

We first give some fairly general lower and upper bounds on $q(n,j,l)$ that, in particular, disprove the conjecture of Erdős, Reid, Schelp and Staton [3]. Both bounds are established by probabilistic methods. After that, we give some constructions, which give upper bounds for $q(n,j,l)$ over a somewhat different range. Finally, we extend the known range of values over which the conjecture of Erdős, Reid, Schelp and Staton does actually hold.

### 4.1. General bounds

In this section, the following Chernoff bound on the tail of the binomial distribution will be used; a proof can be found in [5, Lemma 2].

**Proposition 8.** *Let $X$ be a random variable binomially distributed with parameters $(n, p)$. Then, for $h \geqslant 0$,*

$$\Pr\{X \leqslant (p-h)n\} \leqslant e^{-nh^2/2p} \quad \textit{and} \quad \Pr\{X \geqslant (p+h)n\} \leqslant e^{-nh^2/2(p+h)}.$$

We start with a lower bound for $q(n,j,l)$. As in Section 3, $d_A(v)$ will denote the number of neighbours of the vertex $v$ in the set $A$.

**Theorem 9.** *Let $1 \leqslant j = j(n)$, let $0 \leqslant l = l(n)$ and let $n \to \infty$. If $l = o(n(n+j)/\log(n+j))$, then*

$$q(n,j,l) \geqslant (1+o(1))(l+1)n\left(1+\frac{l}{2n+2j}\right).$$

**Proof.** Let $G \in \mathcal{Q}(n,j,l)$ have size $q(n,j,l)$. We apply a greedy algorithm to $G$. Let $A = \emptyset$ and $B = V(G)$. As long as $\Delta(G[B]) \geqslant n$, move to $A$ a vertex of $G[B]$ of maximum degree. We perform the step at least $l+1$ times, because $\Delta(B) < n$ when we terminate. Hence, for any $n, j, l$, we have $q(n,j,l) \geqslant (l+1)n$, which proves the lemma if $l = o(n+j)$.

So, suppose $l \neq o(n+j)$. Choose some small constant $\eta > 0$. We shall show that, if $n$ is large, then $\Delta(G[B]) \geqslant (1-\eta)n(n+j+l-|A|)/(n+j)$ as long as $\eta l \leqslant |A| \leqslant (1-\eta)l$. Therefore $G$ has at least

$$(1-\eta)\frac{n}{n+j}\sum_{i=\eta l}^{(1-\eta)l}(n+j+l-i) \geqslant (1-2\eta)^2 nl\left(1+\frac{l}{2n+2j}\right)$$

edges. The proof is completed by making $\eta$ small.

To prove the assertion, let $|A| = (1-x)l$ where $\eta \leqslant x \leqslant 1-\eta$, let $b = |B| = n+j+xl$, and let $\varepsilon = \sqrt{x/l}$. Choose a random set $Y \subset B$ by placing each vertex of $B$ into $Y$ independently and with probability $p = (x-\varepsilon)l/b$. The expected value of the binomially distributed variable $|Y|$ is $(x-\varepsilon)l$ so, by Proposition 8, $|Y| < (x-\varepsilon/2)l < l-|A|$ holds with probability at least $1 - \exp\{-\varepsilon^2 l/8x\} = 1 - \mathrm{e}^{-1/8}$.

Let $Z = B\setminus Y$, let $y \in B$ and let $d = d_B(y)$. We may assume that $d \leqslant n(n+j+l)/(n+j)$, for otherwise the assertion is proved. Because $l = \mathrm{o}(n(n+j)/\log(n+j))$, this means that $\log|G| = \mathrm{o}(n^2/d)$. Now $d_Z(y)$ is distributed binomially with parameters $(d, q)$ where $q = 1 - p$. Thus Proposition 8 implies $\Pr\{d_Z(y) \geqslant qd + \eta n/2\} \leqslant \exp\{-\eta^2 n^2/8d\} = \mathrm{o}(1/|G|)$.

Therefore there must exist some $Y$ such that $|Y| + |A| < l$ and $d_Z(y) < qd_B(y) + \eta n/2$ for each $y \in B$. Thus $\Delta(G[Z]) \geqslant n$ and $\Delta(G[Z]) < q\Delta(G[B]) + \eta n/2$. So, in view of $l = \mathrm{o}((n+j)^2)$, we obtain

$$\Delta(G[B]) \geqslant \frac{n - \eta n/2}{q} = \frac{bn(1-\eta/2)}{n + j + \sqrt{xl}} \geqslant (1-\eta)\,\frac{bn}{n+j}\,,$$

as claimed. □

Now let us prove an upper bound for $q(n, j, l)$.

**Theorem 10.** *Let $1 \leqslant j = j(n)$, let $0 \leqslant l = l(n)$ and let $n \to \infty$. If $l = \mathrm{o}(n(n+j))$, then*

$$q(n, j, l) \leqslant (1 + \mathrm{o}(1))(l+1)n\left(1 + \frac{l}{2n+2j}\right).$$

**Proof.** The consideration of $P_{l+1,n}$ together with $j - 1$ extra isolated vertices proves the claim if $l = \mathrm{o}(n)$ or if $j = \mathrm{o}(n)$, so we may assume throughout that $l = \Omega(n)$ and $j = \Omega(n)$.

Suppose first that $l = \Theta(j)$. Fix $0 < \varepsilon < 1$ so that $m = \lceil \varepsilon l \rceil < (n+j)/2$ and $p = (1+\varepsilon)n/(n+j) < 1$. Let $G = P_{l+m,n+j-m}$ and let $V(G) = A \cup B$, where $|A| = l + m, |B| = n + j - m$ and $G[A]$ is complete. Let $H$ be a random spanning subgraph of $G$ formed by choosing each edge independently with probability $p$. By Proposition 8, $\Pr\{e(H) \geq (1+\varepsilon)pe(G)\} \leqslant \exp\{-e(G)\varepsilon^2 p/2(1+\varepsilon)\} = \mathrm{o}(1)$. So almost every subgraph $H$ satisfies $e(H) \leqslant (1+\varepsilon)pe(G)$. Since $\varepsilon$ can be arbitrarily small, we need only show that $H$ is almost certainly in $\mathscr{Q}(n, j, l)$.

Fix an $l$-set $L \subset A \cup B$. Let $x = \lfloor \varepsilon \min\{l/2, (n+j)/4\} \rfloor$. Because $|A\setminus L| \geqslant \varepsilon l > x$, we can choose $X \subset A\setminus L$ with $|X| = x$. Let us estimate from the above the probability that each vertex $y \in X$ has fewer than $n$ neighbours in $Z = (A \cup B)\setminus(L \cup X)$. Now $\Pr\{d_Z(y) < n\} \leqslant \Pr\{d_Z(y) < p|Z| - \varepsilon n/2\}$. Since $d_Z(y)$ is binomially distributed with parameters $(|Z| = n+j-x, p)$, Proposition 8 tells us that the probability of interest is at most $\exp\{-\varepsilon^2 n^2/8p|Z|\} \leqslant \exp\{-\varepsilon^2 n/16\}$. Now the random variables $d_Z(y), y \in X$, are clearly independent, so the probability that $\Delta(G - L) < n$ is at most $\exp\{-\varepsilon^2 nx/16\}$. As there are at most $2^{n+j+l}$ different choices of $L$, we conclude that almost surely $H \in \mathscr{Q}(n, j, l)$, as required.

Suppose now that $l = \mathrm{o}(j)$. Let $C$ be a large constant, take the construction giving an upper bound for $q(n, Cl, l)$ and add $j - Cl$ isolated vertices. Letting $C$ become arbitrarily large proves the lemma in this case also.

It remains to consider the case $l \neq \mathrm{O}(j)$ but $l = \mathrm{o}(n(n+j))$. Let $\varepsilon = 4(l/(nj))^{1/3}$ and $p = (1+\varepsilon)n/(n+j)$. Note that $\varepsilon = \mathrm{o}(1)$; in particular we can assume $p < 1$. Let $H$ be a subgraph of $G = K_{n+j+l}$, choosing edges with probability $p$. Once again, Proposition 8 shows that for any fixed $c > 0$ we have $e(H) \leqslant (1+c)pe(G)$ almost surely. Thus it is enough to show that almost every $H$ belongs to $\mathscr{Q}(n,j,l)$.

Fix an $l$-set $L \subset V(G)$. Take some $X \subset V(G) \backslash L$ of size $x = \lceil \varepsilon j/3 \rceil$. Let $Z = V(G) \backslash (X \cup L)$. Let $y \in X$. Then $\Pr\{d_Z(y) < n\} \leqslant \Pr\{d_Z(y) < p|Z| - \varepsilon n/2\}$, which by Proposition 8 is at most $\exp\{-\varepsilon^2 n^2/8p|Z|\} \leqslant \exp\{-\varepsilon^2 n/16\}$. Hence the probability that each $y \in X$ sends fewer than $n$ edges to $Z$ is at most $\exp\{-\varepsilon^2 xn/16\}$. As there are at most $2^{n+j+l}$ choices of $L$, we conclude that almost surely $H \in \mathscr{Q}(n,j,l)$. $\square$

Now, putting together the bounds of Theorems 9 and 10, we obtain in particular Eq. (2) claimed in the introduction.

Here is a straightforward application of Theorem 9.

**Corollary 11.** *Fix* $0 < c \leqslant 1$. *The minimum size of a graph* $G$ *in which each set of size at least* $c|G|$ *spans a subgraph of maximum degree at least* $n$ *is* $(1+\mathrm{o}(1))[(1-c^2)/2c^2]n^2$ as $n \to \infty$.

**Proof.** The graph $P_{m,n}$ with $m = \lfloor (1-c)n/c \rfloor + 1$ establishes the upper bound. On the other hand, any such graph $G$ belongs to $\mathscr{Q}(n, \lceil cv \rceil - n, v - \lceil cv \rceil)$, where $v = |G|$. By Theorem 9, $e(G) \geqslant (1+\mathrm{o}(1))(1-c^2)nv/2c$. Now the result follows, since $v > n/c$. $\square$

An important instance on its own is the case $l = n$ because, as pointed out in the Introduction, any $G \in \mathscr{B}(n,n)$ belongs to $\mathscr{Q}(n, |G| - 2n, n)$. On the other hand, if we add to $G \in \mathscr{Q}(n,j,n-1)$ a new vertex $x$ connected to everything else, we obtain a $\mathscr{B}(n,n)$-graph. Hence, by (2) we obtain the following result.

**Theorem 12.** *The minimum size of* $G \in \mathscr{B}(n,n)$ *of given order* $v > 2n$ *is*

$$(1+\mathrm{o}(1))\, n^2 \left(1 + \frac{n}{2v-2n}\right) \quad \text{as } n \to \infty.$$

### 4.2. Explicit constructions

The upper bound in Theorem 10 was established by proving probabilistically the existence of a suitable graph. In the case when $j/n$ tends to a fixed integer from above there is a simple explicit construction achieving the same bound. For suppose that $0 < j - n(r-1) = \mathrm{o}(n)$, where $r$ is a natural number.

Let $l + 1 = l_1 + \cdots + l_r$, be a partition into almost equal summands. Let $G$ be the disjoint union of $P_{l_i,n}$, $1 \leqslant i \leqslant r$, plus isolated vertices to ensure that $|G| = n + j + l$. (Note that $j/n > r - 1$ implies that $(l+1) + rn \leqslant n + j + l$.) Now, if we remove at most $l$ vertices, then there is some $i$ such that we remove fewer than $l_i$

vertices from the $i$th component, and the remainer of this component clearly contains a vertex of degree at least $n$. A trivial computation shows that $G$ has the claimed size.

This example gives no information in the range of interest to Erdős, Reid, Schelp and Staton namely $n \geqslant j \geqslant 1$, but here is another construction that does, provided $n < (j-1)l$.

Let $G(n,l)$ be the graph defined as follows. Let $n = lq + r$ with $0 \leqslant r < l$. Take a complete graph on vertex set $A = \{a_1, \ldots, a_{l+1}\}$ and remove the edges of the Hamiltonian path $a_1 a_2 \ldots a_{l+1}$. Take a disjoint set $R$ of $r$ vertices and join everything in $A$ to everything in $R$. Take a new vertex $y$ and join it to $\{a_2, \ldots, a_l\}$. Finally, take disjoint sets $Q_1, \ldots, Q_{l+1}$, each of size $q$, and for $1 \leqslant i \leqslant l+1$, join $v_i$ to everything in $Q_h$ for $h \neq i$. Note that $G(n,l)$ has $n + l + 2 + q$ vertices, and that every vertex in $A$ has degree exactly $n + l - 1$. The size of $G$ is smaller by one than the expression in (1).

**Theorem 13.** *The equation in* (1) *fails to hold if* $n < (j-1)l$.

**Proof.** If $n < (j-1)l$ then $j \geqslant q+2$. So the theorem will be proved if we show that $G(n,l)$, together with $j-q-2$ isolated vertices, is in $\mathcal{Q}(n,j,l)$. Suppose to the contrary that we can remove some set $L$ of size $l$ so that the remaining graph has maximum degree less than $n$. Since $|A| > l$ we have $A \backslash L \neq \emptyset$. Each vertex $x \in A \backslash L$ has degree less than $n$ in $G(n,l) - L$, so it must be connected to every vertex in $L$. Therefore $A \cap L = \emptyset$; for otherwise, we could choose $x \in A \backslash L$ and $y \in A \cap L$ such that $x$ and $y$ are consecutive vertices on the path $a_1 \ldots a_{l+1}$, in which case $xy$ would not be an edge of $G(n,l)$. But the set of vertices connected (in $G(n,l)$) to everything in $A$ is precisely $R$. Therefore $L \subset R$, which implies $|L| \leqslant r < l$, a contradiction. □

### 4.3. The conjecture for small $j$ and $l$

We have shown that the equation in (1) fails to hold for $n < (j-1)l$, though it does hold for $n \geqslant \max(jl, \binom{l+2}{2})$. In fact we can show it to hold in another range, which gives better information if $j$ is smaller than about $l/2$. The proof is a strengthened version of that of Erdős, Reid, Schelp and Staton [3].

**Theorem 14.** *The equation in* (1) *holds provided*

$$n \geqslant \left(j + \frac{1}{2}\right) l + \frac{2j + l}{4j - 2}.$$

**Proof.** The case $j = 1$ of (1) was proved in [3], so assume that $j \geqslant 2$. We prove the theorem by induction on $l$; it is true for $l = 0$, since clearly $q(n,j,0) = n$. Let $l \geqslant 1$ and let $G \in \mathcal{Q}(n,j,l)$. If $\Delta(G) \geqslant n + l$ then the theorem follows by induction, because removing a vertex from $G$ we obtain a graph in $\mathcal{Q}(n,j,l-1)$. So we complete the proof by deriving a contradiction from the two assumptions that $e(G) < (l+1)n + \binom{l+1}{2}$ and $\Delta(G) < n + l$.

Let $H=\{x\in V(G)\colon d(x)\geqslant n\}$ and $h=|H|$. Let us show that $h$ is not large by applying the following algorithm to $G$.

We begin with $A=C=\emptyset$ and $B=V(G)$. Now, choose some vertex $x\in B$ having at least $n$ neighbours in $B$ (if there is one), move it to $A$, and then perform the *n-check*: that is, move to $C$ all $y\in B\cap H$ with $d_{B\cup C}(y)<n$. In fact, for every such $y$ we have $d_{B\cup C}(y)=n-1$. Repeat this procedure as long as possible or until $|A|=l+1$.

Fix the sets $A,B,C$ at the moment the algorithm halts, and let their sizes be $a,b,c$. If $a<l+1$, there must be a vertex $y_1$ in $B\cup C$ that has at least $n$ neighbours in $B\cup C$. Indeed, we may proceed to find a set $Y=\{y_1,\ldots,y_{l+1-a}\}\subset B\cup C$ such that each $y_i$ has at least $n$ neighbours in $(B\cup C)\backslash\{y_1,\ldots,y_{i-1}\}$. As each $y\in C$ has fewer than $n$ neighbours in $B\cup C$, we conclude that $Y\subset B$. Let $R=(B\setminus Y)\cap H$ and $s=|C\cup R|$. Each $x\in R$ has at least $n$ neighbours in $C\cup B$ for otherwise it would belong to $C$. Note that $H=A\cup Y\cup R\cup C$, so $h=l+1+s$.

Counting the number of edges encountered in our investigation we obtain that

$$e(G)\geqslant an+n|Y|+s(n-1)-\binom{s}{2}-s|Y|$$

and also, counting slightly differently the edges incident with $R$,

$$e(G)\geqslant an+n|Y|+\frac{s(n-1)}{2}-\frac{s|Y|}{2},$$

which is better if $s$ is large.

Since $a+|Y|=l+1$ and $|Y|\leqslant l$, these inequalities imply

$$\binom{l+1}{2}>s\left(n-l-\frac{s+1}{2}\right)\quad\text{and}\quad\binom{l+1}{2}>s\left(\frac{n-1}{2}-\frac{l}{2}\right).\qquad(6)$$

The first inequality in (6), which is quadratic in $s$, means that $s$ cannot lie between $n-l-(r+1)/2$ and $n-l+(r-1)/2$, where $r=(4n^2-4n(2l+1)+1)^{1/2}$. Now the assumption of the theorem implies that $l\leqslant 3n/8$ and $n\geqslant 4$, which in turn imply that $r\geqslant n-3$. Suppose that $s\geqslant n-l+(r-1)/2$. Then $s\geqslant n-l+(n-4)/2\geqslant 9n/8-2$. From this, and $l\leqslant 3n/8$, and the second inequality in (6), we obtain $9n^2+32<44n$, which cannot be satisfied for $n\geqslant 4$.

Consequently $s\leqslant n-l-(r+1)/2$, so $h=l+1+s\leqslant n-(r-1)/2$. This implies $jh\leqslant n+j$: to verify this, it is enough to check that $rj\geqslant 2(j-1)n-j$ which, by squaring, is equivalent to $n(2j-1)\geqslant 2j^2l+j$, the assumption of the theorem. Now $\Delta(G)<n+l$, so for every $x\in H$ we may choose a set $D_x$ of $j$ non-neighbours of $x$. Let $D=\cup_{x\in H}D_x$. Then $|D|\leqslant jh\leqslant n+j$. Add to $D$ any further $n+j-jh$ vertices. Since $G\in\mathcal{Q}(n,j,l)$, some $x\in D\cap H$ has at least $n$ neighbours in $D$. But this contradicts the fact that $D$ has at least $j$ non-neighbours of $x$, and the proof is complete. $\square$

## References

[1] B. Bollobás, Random Graphs, Academic Press, New York, 1985, p. xvi+447.

[2] P. Erdős, Problems and results in graph theory, in: G. Chartrand (Ed.), The Theory and Applications of Graphs, Wiley, New York, 1981, pp. 331–341.
[3] P. Erdős, T.J. Reid, R. Schelp, W. Staton, Sizes of graphs with induced subgraphs of large maximum degree, Discrete Math. 158 (1996) 283–286.
[4] P. Erdős, R. Faudree, Restricted size Ramsey numbers for cycles and stars, in: Y. Alavi et al. (Eds.), Proceedings of the Eighth International Conference on Graph Theory, Combinatorics, Algorithms and Applications, New Issues Press, 1998, pp. 353–367.
[5] R. Häggkvist, A. Thomason, Oriented hamilton cycles in digraphs, J. Graph Theory 19 (1995) 471–479.
[6] O. Pikhurko, Size Ramsey numbers of stars versus 3-chromatic graphs, Combinatorica 21 (2001) 403–412.