Discrete Mathematics 233 (2001) 163–174
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

Sparse sets in the complements of graphs with given girth

A.V. Kostochka[a,*,1], D.R. Woodall[b]

[a] Siberian Branch of RAN, Sobolev Institute of Mathematics, Novosibirsk 630090, Russia
[b] School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK

Abstract

A set of edges in a graph is *sparse* if no two of these edges belong to the same clique. It is shown here that if a graph has girth at least 5, 6 or 8 then the largest number of edges in a sparse set in its complement is at most 8, 7 or 6, respectively; this result is complete and best possible. It follows that if $\varepsilon > 0$, then for sufficiently large n there exists a graph with n vertices and chromatic number greater than $n^{1/3-\varepsilon}$, $n^{1/4-\varepsilon}$ or $n^{1/6-\varepsilon}$ whose complement contains no sparse set with more than 8, 7 or 6 edges, respectively. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Sparse set; Girth; Clique covering number; Chromatic number; Complement of a graph

1. Introduction

A set of edges in a graph G is called *sparse* if no two of these edges belong to the same clique in G. Let $h(G)$ denote the cardinality of a largest sparse set of edges in G. This parameter is a lower bound for the *clique covering number* $\mathrm{cc}(G)$, which is the smallest number of cliques in G covering all its edges. Equivalently (and interestingly), $\mathrm{cc}(G)$ is the minimum cardinality of a set M such that G is the intersection graph of a family $\mathscr{G}$ of subsets of M (where subsets can belong to $\mathscr{G}$ with different multiplicities).

Let $\alpha(G)$, $\chi(G)$, $g(G)$ and $\bar{G}$ denote, respectively, the independence number, chromatic number, girth and complement of G. If G has no isolated vertices then it is not difficult to see that

$$\alpha(G) \leqslant h(G) \leqslant \mathrm{cc}(G) \quad \text{and} \quad \alpha(G) \leqslant \chi(\bar{G}) \leqslant \mathrm{cc}(G).$$

Choudum and Parthasarathy [3] conjectured that $\chi(\bar{G}) \leqslant h(G)$ for every graph G without isolated vertices, so that $\alpha(G) \leqslant \chi(\bar{G}) \leqslant h(G) \leqslant \mathrm{cc}(G)$. But Erdős [5] showed that this

* Corresponding author.
E-mail address: sasha@math.nsc.ru (A.V. Kostochka).

[1] This work was carried out while the first author was visiting Nottingham, funded by Visiting Fellowship Research Grant GR/L54585 from the Engineering and Physical Sciences Research Council. The work of this author was also partly supported by grants 96-01-01614 and 97-01-01075 of the Russian Foundation for Fundamental Research.

PII: S0012-365X(00)00235-1

fails for almost all graphs on n vertices. Moreover he proved that, if h is large enough, then

$$\begin{array}{l}\text{there exists a constant } c_h > 0 \text{ such that, for every sufficiently large } n,\\ \text{there is an } n\text{-vertex graph } G \text{ with no isolated vertices}\\ \text{for which } h(G) \leqslant h \text{ and } \chi(\bar{G}) \geqslant n^{c_h}.\end{array} \tag{1.1}$$

Let h_0 be the minimum integer such that (1.1) holds whenever $h \geqslant h_0$, and let h_1 be the maximum integer such that $\chi(\bar{G}) \leqslant h(G)$ whenever $h(G) \leqslant h_1$ and G has no isolated vertices. Erdős [5,4] asked the following questions:

(i) What is the value of h_0?
(ii) What is the value of h_1?
(iii) For $h \geqslant h_0$, how fast can c_h grow with the growth of h?

Brigham and Dutton [1] proved that $h_1 \geqslant 2$. It was proved in [7] that $h_1 = 5$ and $h_0 = 6$ and in [8] that

$$1 - 2\sqrt{2}/(\sqrt{h} - 2) < c_h \leqslant 1 - 1/\sqrt{h} \tag{1.2}$$

for every $h > 6$. (Strictly speaking, these bounds are on the supremum of possible values of c_h.) Furthermore, it was proved in [7] that $h(\bar{G}) \leqslant 6$ for every graph with girth at least 8; for small h, this implies a better lower bound on c_h than the 1 in (1.2).

Motivated by this result and the question (iii), we here determine how large $h(\bar{G})$ can be, subject to weaker lower bounds on the girth $g(G)$ of G. Let $h_{\max}(g)$ denote the maximum value of $h(\bar{G})$ over all graphs with girth at least g. By the above, $h_{\max}(g) \leqslant 6$ if $g \geqslant 8$. There are graphs G with $g(G) = 4$ and $h(\bar{G})$ arbitrarily large, since $K_{p,p}$ minus a 1-factor has girth 4, and the p edges of the removed 1-factor form a sparse set in $\bar{G}$; hence $h_{\max}(g)$ is undefined if $g \leqslant 4$. In Fig. 1 we see examples of graphs with $g(G) = 5$ and $h(\bar{G}) = 8$, with $g(G) = 7$ and $h(\bar{G}) = 7$, and with $g(G) = \infty$ and $h(\bar{G}) = 6$. (The single edges in Fig. 1 are the edges of G, and the double edges form a sparse set in $\bar{G}$.) In view of these examples, the following theorem is complete and best possible, and Corollary 1.2 follows immediately from it.

Theorem 1.1. *If* $g(G) \geqslant 5$ *then* $h(\bar{G}) \leqslant 8$, *if* $g(G) \geqslant 6$ *then* $h(\bar{G}) \leqslant 7$, *and if* $g(G) \geqslant 8$ *then* $h(\bar{G}) \leqslant 6$.

Corollary 1.2. $h_{\max}(5) = 8$, $h_{\max}(6) = h_{\max}(7) = 7$ *and* $h_{\max}(g) = 6$ *if* $g \geqslant 8$.

By a recent result of Krivelevich [9], Theorem 1.1 also implies the following lower bounds on c_h.

Corollary 1.3. *For every* $\varepsilon > 0$, *we can take* $c_8 > \frac{1}{3} - \varepsilon$, $c_7 > \frac{1}{4} - \varepsilon$ *and* $c_6 > \frac{1}{6} - \varepsilon$.

We shall prove Theorem 1.1 in Section 3; the result for $g(G) \geqslant 8$ was proved in [7], but we include a proof here for completeness. In Section 4 we shall prove Corollary 1.3. First, in Section 2, we prove a basic lemma that is needed in Section 3.

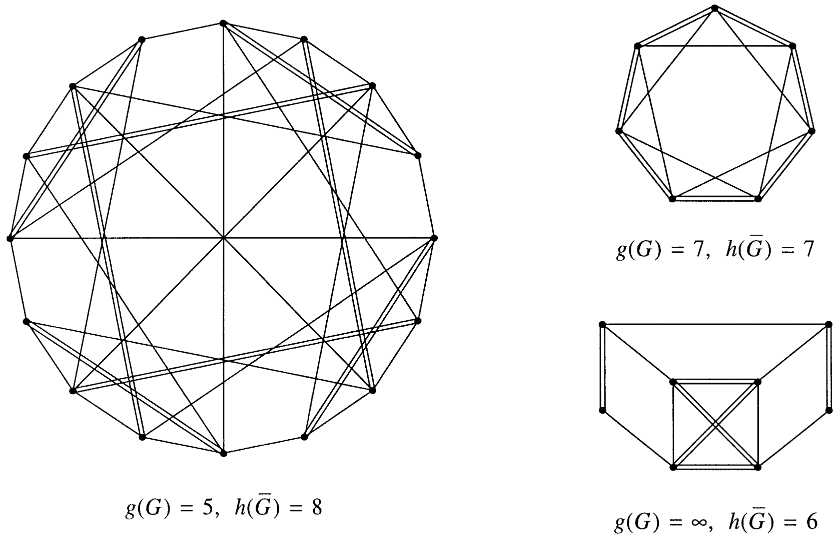

Fig. 1. Graphs to illustrate equality in Theorem 1.1.

Table 1

$n=$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
$m\geqslant$	3	4	5	6	7 ∣	9	10	11 ∣	13	14	15 ∣	17					$\Rightarrow$ girth $\leqslant 7$
$m\geqslant$	3	4	5 ∣	7	8 ∣	10	11 ∣	13	15	17	19 ∣	22	23	25 ∣			$\Rightarrow$ girth $\leqslant 5$
$m\geqslant$	3	4 ∣	6	7 ∣	9 ∣	11	13 ∣	16	17	19 ∣	22	25	28	31 ∣	34	36	$\Rightarrow$ girth $\leqslant 4$

2. A basic lemma

Before starting on the proof proper of Theorem 1.1, it is convenient to tabulate the numbers of edges that are required to ensure that a graph with n vertices has a cycle of length (at most) 4, 5 or 7. We do this in Table 1, with proof in Lemma 2.1. (The vertical lines in the table indicate places where the method of proof changes.) The figures in this table are adequate for our purposes; we have made a reasonable effort to get good bounds, but the reader should not assume that these figures are necessarily best possible.

Lemma 2.1. *If G is a graph with n vertices and m edges, then G has girth at most* 7, 5 *or* 4 *if m is at least as large as is specified in Table* 1 *for the relevant values of n and the girth.*

Proof. Let $g=7$, 5 or 4 and let G be a minimal counterexample to the lemma. Let G have n vertices and m edges, so that m is exactly equal to the lower bound specified for m in Table 1, and $g(G)\geqslant g+1$. Note that G is connected and has no cut-edge, since identifying two vertices in different components, or contracting a cut-edge, would reduce n by 1 and reduce m by at most 1 without creating any shorter cycles, and so

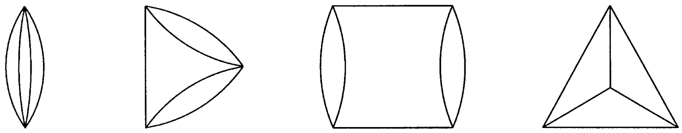

Fig. 2. Multigraphs with two more edges than vertices.

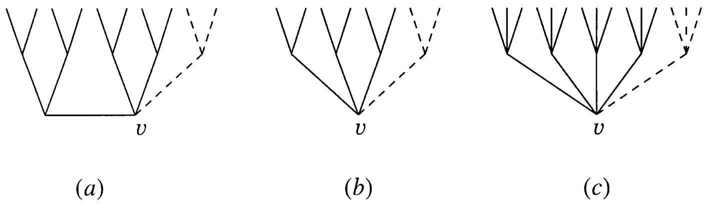

Fig. 3. Subtrees of graphs with large girth and minimum degree.

would produce a smaller counterexample than G. The minimum degree of G is denoted by $\delta(G)$; clearly $\delta(G) \geqslant 2$.

If $m = n$, then G is a cycle, and so $n \geqslant g + 1$, contrary to the figures in Table 1. If $m = n + 1$, then G is a θ-graph (two vertices of degree 3 connected by three paths), and it is easy to see that for $g(G) > 7$, 5 or 4 we require $n \geqslant 11$, 8 or 7, respectively, again contradicting Table 1. If $m = n + 2$, then G has at most four vertices with degree 3 or more, and on suppressing vertices of degree 2 we obtain a multigraph H which is easily seen to be one of the four in Fig. 2. Then G is obtained from H by inserting vertices of degree 2 into some of the edges of H, and it is not difficult to see that for $g(G) > 7$, 5 or 4 we require $n \geqslant 14$, 10 or 8, contrary to Table 1. (The minimum value of n is obtained always when $H = K_4$, and also, in two cases, when H is the sheaf of four parallel edges.)

If $(n, m) = (14, 17)$, then G must contain a vertex u with degree 2; then $G - u$ has 13 vertices and 15 edges and so $g(G - u) \leqslant 7$. This completes the proof for $g = 7$. For $g = 5, 4$ and larger values of n, the lower bound given for m in Table 1 increases by at least 2 for every increase of 1 in n, and so we may assume that $\delta(G) \geqslant 3$, except in the cases $(n, m, g) = (15, 23, 5)$ and $(11, 17, 4)$.

Suppose first that $g = 5$. Let v be an arbitrary vertex of G. Then G contains the tree with 14 distinct vertices depicted by the unbroken edges in Fig. 3(a). This proves the result for $n \leqslant 13$, and it also shows that if $(n, m) = (14, 21)$ then G has diameter at most 3. If $(n, m) = (15, 23)$ and G contains a vertex u with degree 2, it follows that $G - u$ has diameter at most 3 and so $g(G) \leqslant 5$. If G has no vertex with degree 2 and $(n, m) = (14, 22)$, $(15, 23)$ or $(16, 25)$, then G has a vertex v with degree at least 4, and we see from Fig. 3(a) that $n \geqslant 17$, a contradiction. This completes the proof for $g = 5$.

Finally, suppose that $g = 4$. Since $\delta(G) \geqslant 3$ except when $(n, m) = (11, 17)$, G must contain the tree represented by the unbroken lines in Fig. 3(b), so that $n \geqslant 10$. Moreover, if $(n, m) = (10, 15)$ then G has diameter at most 2. If $(n, m) = (11, 17)$ and G contains a vertex u with degree 2, then $G-u$ has diameter at most 2 and so $g(G)\leqslant 4$. If G has no vertex with degree 2 and $(n,m)=(10,16)$, $(11,17)$ or $(12,19)$, then G has a vertex v with degree at least 4, and we see from Fig. 3(b) that $n\geqslant 13$, a contradiction.

For $n\geqslant 12$, the lower bound given for m in Table 1 increases by at least 3 for every increase of 1 in n, and so we may assume that $\delta(G)\geqslant 4$, except in the case $(n,m)=(18,36)$. It follows from the tree represented by the unbroken lines in Fig. 3(c) that $n\geqslant 17$. For $(n,m)=(17,34)$ or $(18,36)$ it is well known [2,6,10,11] that there is no 4-regular graph with girth 5. This disposes of $n=17$, but for $n=18$ we still have to consider the possibility that $\delta(G)=3$. By the tree argument of Fig. 3(c), G contains no vertex v with degree 6, and if G contains a vertex v with degree 5 then v has at least three neighbours x,y,z of degree 3 (otherwise $n\geqslant 19$). But then $G-\{v,x,y,z\}$ has $18-4=14$ vertices and $36-11=25$ edges and so $g(G)\leqslant 4$. This completes the proof of Lemma 2.1. □

3. Proof of the main theorem

We now start the proof proper of Theorem 1.1. Let $(g,h)=(7,7)$, $(5,8)$ or $(4,9)$. Let $G=(V,E)$ be a graph with $g(G)\geqslant g+1$ and $h(\bar{G})=h$, and let $\tilde{G}=(V,\tilde{E})$ where $\tilde{E}$ is a sparse set of h edges in $\bar{G}$. We seek a contradiction. We assume that every vertex in V is incident with an edge of $\tilde{E}$ (otherwise we can delete it from both G and $\tilde{G}$). We shall say that two edges in $\tilde{E}$ are *disjoint* if they have no vertex in common, and *adjacent* if there is an edge of G joining two of their endvertices (as there must be, since $\tilde{E}$ is sparse in $\bar{G}$). The term *adjacent* will always refer to adjacency in G, not in $\bar{G}$ or $\tilde{G}$. The set of neighbours of a vertex a in G will be denoted by $N_G(a)$. Let $n:=|V|$ and $m:=|E|$. We start by establishing two principles that will be used frequently throughout the proof.

Principle 1. If $\tilde{G}$ contains edges ab, bc, de, fg, where distinct letters denote distinct vertices, then b is adjacent in G to at least one of the vertices d,e,f,g.

Proof. Suppose not. Since $\tilde{E}$ is sparse in $\bar{G}$, each two edges of $\tilde{E}$ are adjacent in G. Moreover, $g(G)\geqslant 5$. Thus $ac\in E$, and after interchanging d with e and/or f with g if necessary, $ad,ce,af,cg\in E$. But there is at least one edge of G between de and fg, and this must create a 3-cycle or a 4-cycle, a contradiction. □

Principle 2. If $\tilde{G}$ contains edges ab, bc, de, ef, where distinct letters denote distinct vertices, then either $be\in E$ or $\{a,b,c\}$ is joined to $\{d,e,f\}$ by a matching of three nonadjacent edges in G, and the only other edges of G connecting these six vertices are ac and df.

Proof. Suppose $be\notin E$. Since $ac,df\in E$, no vertex can be adjacent in G to a and c or to d and f. In particular, and since $g(G)\leqslant 5$, G contains at most one of the edges

ad, cd, af, cf. If it contains none of them, then we may suppose (after interchanging a with c and/or d with f if necessary) that $ae, bd \in E$, and then there can be no edge of G between $\{b,c\}$ and $\{e,f\}$. But if G contains one of the edges ad, cd, af, cf, say ad, then it contains bf or ce; if bf, then it does not contain bd, cd or be and so must contain ce; if ce, then it does not contain ae, af or be and so must contain bf. Thus we have the required matching $\{ad, bf, ce\}$. It is easy to see that there is no other edge of G between $\{a,b,c\}$ and $\{d,e,f\}$. □

In the rest of the proof, the term 3-*matching* will be used exclusively to denote a matching of the type that must exist by Principle 2.

Lemma 3.1. *$\tilde{G}$ is a disjoint union of paths.*

Proof. $\tilde{G}$ has maximum degree at most 2, since if $ab, ac, ad \in \tilde{E}$ then b, c, d would form a triangle in G (otherwise $\tilde{E}$ would not be sparse in $\bar{G}$). Thus, it suffices to show that $\tilde{G}$ contains no cycle. Suppose $\tilde{G}$ contains a cycle $\tilde{C} = v_1v_2\ldots v_lv_1$. Since $\tilde{E}$ is sparse in $\bar{G}$, all the edges v_iv_{i+2} (subscripts modulo l) are in G. Now, $l \leqslant |\tilde{E}| \leqslant 9$.

If $l = 9$, then G contains a 9-cycle $v_1v_3v_5v_7v_9v_2v_4v_6v_8v_1$, and it must also contain an edge between $\{v_1, v_2\}$ and $\{v_5, v_6\}$; hence G must contain a 3-cycle or a 4-cycle, which is impossible.

If $l = 8$ or 6 then G contains two 4-cycles or two triangles, neither of which is possible.

If $l = 7$ then G contains a 7-cycle $C = v_1v_3v_5v_7v_2v_4v_6v_1$, which is a contradiction if $(g,h) = (7,7)$. If $(g,h) = (5,8)$ or $(4,9)$, then $\tilde{E}$ contains another edge ab. Since $g(G) \geqslant 5$, it is easy to see that $N_G(a) \cap C \subseteq \{v_i, v_{i+1}\}$ and $N_G(b) \cap C \subseteq \{v_j, v_{j+1}\}$ for some i and j. But then there is no edge of G between ab and some edge v_kv_{k+1} of $\tilde{C}$, a contradiction.

If $l = 5$ then G contains a 5-cycle C, and no other vertex can be adjacent in G to more than one vertex of C, so that $\tilde{E}$ cannot be sparse in $\bar{G}$.

Finally, suppose $l = 4$. If all $h-4$ edges of $\tilde{G} - \tilde{C}$ are mutually nonadjacent, then we can apply Principle 1 to each pair of them, with each of the four 2-paths $v_iv_{i+1}v_{i+2}$ in $\tilde{C}$, in order to deduce that each vertex v_i is adjacent to at least $h-5$ vertices outside $\tilde{C}$. There are at least $\binom{h-4}{2}$ edges of G between the edges of $\tilde{G} - \tilde{C}$, and so, counting the edges v_1v_3 and v_2v_4, G has $2h-4$ vertices and at least $\binom{h-4}{2} + 4(h-5) + 2$ edges. For $h = 7$, 8 or 9, we have $n = 10$, 12 or 14 and $m \geqslant 13$, 20 or 28, easily enough to force a 7-, 5- or 4-cycle, respectively, according to Table 1. In fact, for $h = 8$ or 9 we can conclude by the same argument that $\tilde{G} - \tilde{C}$ cannot contain even $h-5$ nonadjacent edges, since their vertices together with those of $\tilde{C}$ would induce a subgraph H of G with $2h-6$ vertices and at least $\binom{h-5}{2} + 4(h-6) + 2$ edges. For $h = 8$ or 9 this gives $|V(H)| = 10$ or 12 and $|E(H)| \geqslant 13$ or 20, enough to force a 5- or 4-cycle, respectively; thus in these cases $\tilde{G} - \tilde{C}$ has at least two vertices of degree 2.

So we may suppose that outside $\tilde{C}$ there are edges ab, bc of $\tilde{E}$. By Principle 2 applied between the 2-paths abc and $v_iv_{i+1}v_{i+2}$ $(1 \leqslant i \leqslant 4)$, b is adjacent in G to at least two vertices of $\tilde{C}$, necessarily consecutive, say v_1 and v_2, and then (after interchanging a with c if necessary) $av_3, cv_4 \in E$. Thus, G contains the 7-cycle $C = av_3v_1bv_2v_4ca$. This is a contradiction if $(g,h) = (7,7)$.

If $(g,h) = (5,8)$ or $(4,9)$, then we have seen that $\tilde{G} - \tilde{C}$ has a vertex $e \neq b$ with degree 2, and so there is another 7-cycle in G. If $e = a$ or c, then the new 7-cycle includes a chord of C and so $g(G) \leqslant 4$, a contradiction. Thus $\tilde{G} - \tilde{C}$ contains edges ab, bc, de, ef, where distinct letters represent distinct vertices. Since b is nonadjacent to some edge of $\tilde{C}$, Principle 1 implies that b is adjacent to de and ef, which (since $df \in E$ and so bd, bf are not both in E) forces $be \in E$. But b and e are each adjacent to two vertices of $\tilde{C}$, and so G contains a 3-cycle or a 4-cycle. This contradiction completes the proof of Lemma 3.1. □

Lemma 3.1 will be assumed implicitly throughout the whole of the rest of the proof of Theorem 1.1.

Lemma 3.2. *$\tilde{G}$ cannot contain edges ab, bc, cd, ef, fg, gh, where distinct letters represent distinct vertices.*

Proof. Suppose it does. Then $ac, bd, eg, fh \in E$. Let C denote the 4-cycle $bfcgb$. Clearly $|E(G \cap C)| \leqslant 3$, and $|E(G \cap C)| \geqslant 1$ since otherwise there is no edge of G between bc and fg. Suppose that $E(G \cap C)$ contains bf but not cg. Then the 3-matching between $\{b,c,d\}$ and $\{f,g,h\}$ must contain edges ch and dg, so that $ah, bg, de, cf, df, bh \notin E$ (because $ac, bd, eg, fh \in E$). Thus, the edge of G joining $\{c,d\}$ with $\{e,f\}$ must be ce and that joining $\{a,b\}$ with $\{g,h\}$ must be ag, giving a 4-cycle $ageca$. This contradiction shows that $E(G \cap C) = \{bf, cg\}$ or (reversing the labelling of one path) $\{bg, cf\}$: suppose the latter. Then it is easy to see from Principle 2 that $ae, dh \in E$, so that G contains the 8-cycle $aegbdhfca$.

If any of a, d, e, h, say h, has degree 2 in $\tilde{G}$, then by the same reasoning G contains exactly two nonadjacent edges of the 4-cycle $bgchb$, and so it contains edge ch, giving a triangle $chfc$. This contradiction shows that every other edge of $\tilde{G}$ is disjoint from the vertices a, d, e, h. Let uv be another edge of $\tilde{G}$. For there to be an edge of G between $\{u,v\}$ and each other edge of $\tilde{G}$ without creating a 5-cycle in G, u must be adjacent to b, c and v to f, g or vice versa, thus creating a 6-cycle $ubgvfcu$ or $vbgufcv$. This gives a contradiction if $(g,h) = (7,7)$. If $(g,h) = (5,8)$, then we get the contradiction because there is another such edge wx in $\tilde{E}$ (possibly $w = v$), which causes G to contain a 4-cycle of the form $bucxb$.

So suppose $(g,h) = (4,9)$. Among the three edges of $\tilde{G}$ not mentioned in the statement of the Lemma, there are two that are disjoint, say uv and wx. Since b is not adjacent in G to either vertex of the edge ef of $\tilde{G}$, it follows from Principle 1 that b is adjacent in G to at least one vertex of each of the edges uv and wx. Similarly, so is g. Since $bg \in G$, w.l.o.g. b is adjacent to u and w and g is adjacent to v and x. But there is an edge of G between $\{u,v\}$ and $\{w,x\}$, and this creates a 3-cycle or a 4-cycle. This contradiction completes the proof of Lemma 3.2. □

Lemma 3.3. *$\tilde{G}$ cannot contain edges ab, bc, de, ef, fg, gh, where distinct letters represent distinct vertices.*

Proof. Suppose it does. Then $ac, df, eg, fh \in E$. Let $B:=\{be, bf, bg\}$. The 3-matchings implied by Principle 2 imply that $B \cap E \neq \emptyset$, and clearly $\{be, bg\} \not\subseteq E$. If $B \cap E = \{bf\}$ and we take care to avoid triangles in G, then we quickly find that the 3-matchings lead to a 4-cycle $acega$ or $acgea$. If $B \cap E = \{bg\}$ then the 3-matchings force $af, ce \in E$ (after interchanging a with c if necessary), so that $bd \in E$ and we have a 7-cycle $C_1 = acegbdfa$. The only other (essentially different) possibility is that $B \cap E = \{bf, bg\}$, in which case $ad, ce \in E$ (after interchanging a with c if necessary) and we have a 7-cycle $C_2 = acegbfda$.

Suppose that $\tilde{E}$ contains an edge hw incident with h. The presence of edge bh or cw in G would give a contradiction, since the first implies the existence of a 4-cycle $bhfdb$ or a triangle $bhfb$ (depending on which of C_1, C_2 is present in G), and the latter gives a 4-cycle $cwgec$. So the 3-matching between $\{a, b, c\}$ and $\{g, h, w\}$ must contain edges ch and aw. Now the presence of C_1 in G would imply the existence of a 4-cycle $chfac$, and so it must be C_2 that is present in G. The sparseness of $\tilde{E}$ implies that there is an edge of G between $\{d, e\}$ and $\{h, w\}$; but the existence of edge dh, dw, eh or ew in G would now give a triangle $dhfd$, $dwad$, $ehce$ or $ewge$. This contradiction shows that no edge of $\tilde{E}$ is incident with h.

If $(g, h) = (7, 7)$, we already have a contradiction from the existence of C_1 or C_2. So suppose otherwise, which means that there are at least two further edges of $\tilde{E}$. By Lemma 3.2 and the previous paragraph, we may assume that at least one such edge uv has none of a, c, d, h among its endvertices. For there to be an edge of G between uv and each of edges ab, bc, de, ef, fg, gh of $\tilde{E}$, without creating a 4-cycle, it is not difficult to see that u (say) must be adjacent to e and f and v (say) to b and h. This is true whichever of the two 7-cycles G contains, and in each case it creates a 5-cycle. This gives a contradiction if $(g, h) = (5, 8)$. If $(g, h) = (4, 9)$ then there is another such edge $u'v'$, and even if $\{u, v\} \cap \{u', v'\} \neq \emptyset$ we must have a 4-cycle of the form $eufu'e$ or $bvhv'b$. This contradiction completes the proof of Lemma 3.3. □

Lemma 3.4. *$\tilde{G}$ cannot contain a path of four or more edges.*

Proof. Let $\tilde{P} = v_0 v_1 \ldots v_l$ be a longest path in $\tilde{G}$. The previous two lemmas both imply $l \leqslant 6$, and Lemma 3.3 also implies (if $l \geqslant 4$) that the edges of $\tilde{G} - \tilde{P}$ are all disjoint. There are three cases to consider.

Case 1: $l = 6$. Then G contains edges v_0v_2, v_2v_4, v_4v_6, v_1v_3 and v_3v_5, so that it cannot contain v_0v_6 or v_1v_5 and must contain v_0v_5 and v_1v_6 by Principle 2. Thus there is a 7-cycle $v_0v_2v_4v_6v_1v_3v_5v_0$. Since $g(G) \geqslant 5$, if any other vertex is adjacent to two vertices of $\tilde{P}$, then they must be v_0 and v_6 or v_i and v_{i+1} for some i. For each other edge uv of $\tilde{E}$, we therefore require one of u, v to be adjacent to v_1 and v_2 and the other to be adjacent to v_4 and v_5. If there are two such edges uv then we have a 4-cycle. If there is only one, then $(g, h) = (7, 7)$ and the existence of the 7-cycle already gives a contradiction. This shows that $l \neq 6$.

Case 2: $l = 5$. Then the edge of G between v_0v_1 and v_4v_5 is either v_0v_5 or v_1v_4, so that G contains one of the paths $P_1 = v_1v_3v_5v_0v_2v_4$ and $P_2 = v_0v_2v_4v_1v_3v_5$. For each edge $uv \in E(\tilde{G} - \tilde{P})$, u (say) is adjacent in G to at least two vertices of $\tilde{P}$, so that there are at least 7 edges of G joining the 7 vertices in $V(\tilde{P}) \cup \{u\}$. This is impossible if $(g,h) = (7,7)$.

So suppose $(g,h) = (5,8)$ or $(4,9)$. $\tilde{G} - \tilde{P}$ consists of $h - 5$ nonadjacent edges. By Principle 1, each of v_1, v_2, v_3, v_4 is adjacent to at least $h - 6$ of the $2h - 10$ vertices of these $h-5$ edges. If $v_1v_4 \in E$, then among the $2h-6$ vertices mentioned (not including v_0 and v_5) there are therefore at least $\binom{h-5}{2} + 4(h-6) + 3$ edges. For $(g,h) = (5,8)$ or $(4,9)$ we have 10 or 12 vertices and at least 14 or 21 edges, enough to force a 5-cycle or 4-cycle, respectively, from Table 1. If $v_1v_4 \notin E$, then v_1 is adjacent to neither of v_4, v_5, and so, by Principle 1, it is adjacent to at least $h - 5$ vertices of $\tilde{G} - \tilde{P}$; we obtain the same contradiction.

Case 3: $l = 4$. By Principle 1, each of v_1, v_2 and v_3 is adjacent to at least $h - 5$ of the $2h - 8$ vertices of the $h - 4$ edges in $\tilde{G} - \tilde{P}$. Among the $2h - 5$ vertices mentioned (not including v_0 and v_4) there are therefore at least $\binom{h-4}{2} + 3(h-5) + 1$ edges of G (including the edge v_1v_3). For $(g,h)=(7,7)$, $(5,8)$ or $(4,9)$ we have 9, 11 or 13 vertices and at least 10, 16 and 23 edges, enough to force a 7-, 5- or 4-cycle, respectively, from Table 1. This contradiction completes the proof of Lemma 3.4. □

Lemma 3.5. *$\tilde{G}$ cannot contain edges ab, bc, cd, ef, fg, where distinct letters represent distinct vertices.*

Proof. Suppose it does; let $\tilde{E}'$ be the set of five edges mentioned. Note that, by Lemmas 3.2 and 3.3, every edge in $\tilde{E} \backslash \tilde{E}'$ is disjoint from every edge in $\tilde{E}'$. Let $uv, v'w \in \tilde{E} \setminus \tilde{E}'$, chosen so that $v = v'$ if possible. Let $A := \{bf, cf\}$. There are three cases to consider.

Case 1: $|E \cap A| = 0$. Then the necessary 3-matchings imply that (after interchanging e and g if necessary) there is a 7-cycle $afdbegca$. By Principle 1, since each of b, c, f is nonadjacent in G to some edge in $\tilde{E}'$, each of b, c, f is adjacent in G to one of u, v. But then two are adjacent to the *same* one of u, v, which must create a 4-cycle or (if the two are b, c) a 5-cycle. This is a contradiction if $(g,h) = (7,7)$ or $(5,8)$. If $(g,h) = (4,9)$, then $\tilde{E}$ contains another edge xy disjoint from those already mentioned, and then two of b, c, f must also be adjacent to the same one of x, y, now creating a 4-cycle in each case.

Case 2: $|E \cap A| = 1$, say $E \cap A = \{cf\}$. Then (after interchanging e and g if necessary) the 3-matchings imply the existence of a path $dbegacf$. As in Case 1, b and f are adjacent to every edge of $\tilde{E} \setminus \tilde{E}'$, but now c need not be adjacent to uv. However, if it is not, then (by Principles 1 and 2) c is adjacent to every other edge of $\tilde{E}$. Thus one of v', w must be adjacent to c and f (creating a triangle), or to b and c (creating a 6-cycle) or to b and f (creating a 7-cycle). This is a contradiction if $(g,h) = (7,7)$. If not, then $\tilde{E}$ contains another edge xy disjoint from the seven already mentioned (by

Table 2

	$(g,h)=(7,7)$			$(g,h)=(5,8)$			$(g,h)=(4,9)$		
k	n'	m'		n'	m'		n'	m'	
3	5	4	$(0+3\times 2-2)$	7	8	$(1+3\times 3-2)$	9	13	$(3+3\times 4-2)$
2	8	8	$(3+2\times 3-1)$	10	13	$(6+2\times 4-1)$	12	19	$(10+2\times 5-1)$
1	11	14	$(10+1\times 4-0)$	13	20	$(15+1\times 5-0)$	15	27	$(21+1\times 6-0)$

Lemma 3.2 and the choice of v, v'); thus, G must contain a 4-cycle of the form $cwbxc$ or $bwfxb$ or a 5-cycle of the form $cwbxfc$, which is a contradiction if $(g,h)=(5,8)$. Finally, if $(g,h)=(4,9)$ then $\tilde{E}$ contains yet another edge $y'z$, and (even if $y=y'$, when $xz \in E$) we cannot avoid a 4-cycle of the form mentioned.

Case 3: $A \subseteq E$. If $v=v'$ then v cannot be adjacent to both of b,c, and so we can interchange the roles of e,f,g and u,v,w and get a contradiction by Case 2. Thus, we may assume that all edges of $\tilde{E}\setminus\tilde{E}'$ are disjoint. Thus, $|V\setminus\{a,d,e,g\}|=2h-7$ and the number of edges of G between these vertices is (by Principle 1) at least $\binom{h-5}{2}+3(h-6)+2$. For $h=8$ or 9 we have 9 or 11 vertices and at least 11 or 17 edges, enough to force a 5- or 4-cycle, respectively, from Table 1. Also, if $h=7$, there are at least three edges of G between $\{a,b,c,d,e,f\}$ and each of $\{u,v\}$ and $\{v',w\}$, and one edge between these last two sets, giving (with the path $acfbd$) 10 vertices and 11 edges, enough to force a 7-cycle. This completes the proof of Lemma 3.5. □

Finally, we are in a position to complete the proof of Theorem 1.1 itself. Suppose first that $\tilde{E}$ contains edges ab, bc, cd. Then, by Lemmas 3.4 and 3.5, all remaining edges of $\tilde{E}$ are disjoint from these and from each other. By Principle 1, among the $2h-4$ vertices of $G-\{a,d\}$ there are at least $\binom{h-3}{2}+2(h-4)$ edges. For $h=7$, 8 or 9 this gives 10, 12 or 14 vertices and 12, 18 or 25 edges, enough to force a 7-, 5- or 4-cycle, respectively, according to Table 1.

So we may suppose that each component of $\tilde{G}$ has one or two edges. Let the edges of $\tilde{G}$ be a_ib_i, b_ic_i and d_je_j $(1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant l,\ l=h-2k)$, where distinct symbols denote distinct vertices. Let G' be the subgraph of G induced by the vertices $b_1,\dots,b_k,d_1,\dots,d_l,e_1,\dots,e_l$. We may assume $k<4$, since if $k=4$ then $b_1,\dots,b_4$ induce a subgraph with at least 4 edges, by Principle 1. If $k=0$, then G has $2h$ vertices and at least $\binom{h}{2}$ edges; for $h=7,8,9$ this is $14,16,18$ vertices and $21,28,36$ edges, enough to force a 7-, 5- or 4-cycle, respectively, from Table 1.

So $k=1$, 2 or 3. It follows from Principle 1 that each vertex b_i has degree at least $k+l-2=h-k-2$ in G', since if b_1 is not adjacent to b_2, say, then either it is nonadjacent to both of a_2 and b_2 or it is nonadjacent to both b_2 and c_2. Thus G' has $n'=2h-3k$ vertices and $m' \geqslant \binom{h-2k}{2}+k(h-k-2)-(k-1)$ edges by Principle 1, since the term $h(h-k-2)$ counts twice any edges between the vertices $b_1,\dots,b_k$, of which there are at most $k-1$. We therefore get the values in Table 2.

From Table 1 we see that $g(G) \leqslant g$, a contradiction, with three possible exceptions which we now examine. If $(g,h,k,n',m')=(4,9,1,15,27)$, we note that b_1 has degree at least 6 in G', which implies (by the tree argument of Fig. 3(b), since $n'<19$) that

G' contains a vertex v of degree 2 or less. Then $G' - v$ has 14 vertices and at least 25 edges, enough to force a 4-cycle. If $(g,h,k,n',m')=(7,7,2,8,8)$, we note that, to avoid a 7-cycle, G' must be an 8-cycle, which is not possible since b_1 and b_2 each have degree at least 3 in G'. Finally, if $(g,h,k,n',m')=(7,7,3,5,4)$, suppose w.l.o.g. b_1b_2 and $b_2b_3 \in E$ (since each b_i is adjacent to at least one other b_j). Then the 3-matching between $\{a_1,b_1,c_1\}$ and $\{a_3,b_3,c_3\}$, together with the edges b_1b_2, b_2b_3, a_1c_1 and a_3c_3, gives 7 edges on 7 vertices and so forces a 7-cycle. This finally completes the proof of Theorem 1.1.

4. Proof of Corollary 1.3

In order to state the lemma of Krivelevich that we shall use, we first need some more definitions. If $H=(V,E)$ is a graph with at least three vertices, its *density* $\rho(H)$ is the maximum value of the ratio $(|E(F)|-1)/(|V(F)|-2)$ over all subgraphs F of H with at least three vertices. Given a family $\mathscr{H}=\{H_1,\ldots,H_t\}$ of graphs (each with at least three vertices), let $\rho(\mathscr{H})=\min\{\rho(H_i)\colon 1\leqslant i\leqslant t\}$. As usual, $G(n,p)$ will denote a random graph on n labelled vertices in which each edge is present with probability p (independent of all other edges). Krivelevich's lemma ([9], Lemma 6.1) is as follows.

Lemma 4.1. *Let $\mathscr{H}=\{H_1,\ldots,H_t\}$ be a family of graphs, each with at least three vertices, such that $\rho(\mathscr{H})>1$. Then there exist positive constants $C=C(\mathscr{H})$ and $c=c(\mathscr{H})$ such that, if p is a probability satisfying*

$$\frac{(\ln n)^2}{n}\leqslant p\leqslant cn^{-1/\rho(\mathscr{H})},$$

then, with probability tending to 1 as $n\to\infty$, a random graph $G=G(n,p)$ has the following property: If $\mathscr{G}$ is any family of edge-disjoint subgraphs of G, each isomorphic to a graph in $\mathscr{H}$, then after the deletion of all edges of all graphs in $\mathscr{G}$, the resulting subgraph G_0 on n vertices does not contain an independent set of size $\lceil C\ln n/p\rceil$.

Let $\mathscr{H}=\{C_3,\ldots,C_{k+1}\}$ be the family of cycles of length at most $k+1$. Then $\rho(\mathscr{H})=k/(k-1)>1$. Thus, putting $p=cn^{-(k-1)/k}$ in Lemma 4.1, we find that for sufficiently large n there exists a graph $G_0=G_0(n,k)$ on n vertices with girth at least $k+2$, with no isolated vertices and with independence number $\alpha(G_0(n,k))\leqslant(C\ln n/c)n^{(k-1)/k}$. It follows that, given any $\varepsilon>0$, if n is large enough then

$$\chi(G_0(n,k))\geqslant\frac{n}{\alpha(G_0(n,k))}\geqslant\frac{c}{C\ln n}n^{1/k}>n^{1/k-\varepsilon}.$$

By Theorem 1.1 with $k=g-2$, $h(\overline{G_0(n,3)})\leqslant 8$, $h(\overline{G_0(n,4)})\leqslant 7$ and $h(\overline{G_0(n,6)})\leqslant 6$. It now follows from (1.1) that, given any $\varepsilon>0$, we can take $c_8>\frac{1}{3}-\varepsilon$, $c_7>\frac{1}{4}-\varepsilon$ and $c_6>\frac{1}{6}-\varepsilon$. This completes the proof of Corollary 1.3.

Acknowledgements

The authors are grateful to an unknown referee who checked the details of the proof with great thoroughness and made several helpful suggestions.

References

[1] R.C. Brigham, R.D. Dutton, On clique covers and independence numbers of graphs, Discrete Math. 44 (1983) 139–144.
[2] W.G. Brown, On the non-existence of a type of regular graphs of girth 5, Canad. J. Math. 19 (1967) 644–648.
[3] S.A. Choudum, K.R. Parthasarathy, The edge clique cover number of products of graphs, J. Math. Phys. Sci. 10 (1976) 255–261.
[4] P. Erdős, Problems and results in graph theory, in: The Theory and Applications of Graphs, Wiley, New York, 1981, pp. 331–341.
[5] P. Erdős, On the covering of the vertices of a graph by cliques, J. Math. Res. Exposition 2 (1982) 93–96.
[6] A.J. Hoffman, R.R. Singleton, On Moore graphs with diameters two and three, IBM J. Res. Develop. 4 (1960) 497–504.
[7] A.V. Kostochka, Maximum set of edges no two covered by a clique, Combinatorica 5 (1985) 229–235.
[8] A. V. Kostochka, On the number of graph edges not contained in common cliques, in: Graphs, Hypergraphs and Applications, Teubner-Texte zur Mathematik, Vol. 73, Teubner, Leipzig, 1985, pp. 110–114.
[9] M. Krivelevich, On the minimal number of edges in color-critical graphs, Combinatorica 17 (1997) 401–426.
[10] N. Robertson, The smallest graph of girth 5 and valency 4, Bull. Amer. Math. Soc. 70 (1964) 824–825.
[11] N. Sauer, Extremaleigenschaften regulärer Graphen gegebener Taillenweite, I und II, Sitzungsberichte Österr. Akad. Wiss. Math. Nat. Kl., Abt. II 176 (1967) 9–43.