Available online at www.sciencedirect.com
SCIENCE DIRECT

Journal of Combinatorial Theory, Series B 88 (2003) 329–346

Journal of Combinatorial Theory Series B

http://www.elsevier.com/locate/jctb

Maximum cuts and judicious partitions in graphs without short cycles

Noga Alon,[a,b,1] Béla Bollobás,[c,d,2] Michael Krivelevich,[b,3] and Benny Sudakov[e,f,4]

[a] Institute for Advanced Study, Princeton, NJ 08540, USA
[b] Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel
[c] Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA
[d] Trinity College, Cambridge CB2 1TQ, UK
[e] Department of Mathematics, Princeton University, Princeton, NJ 08540, USA
[f] Institute for Advanced Study, Princeton, NJ 08540, USA

Received 6 March 2002

Abstract

We consider the bipartite cut and the judicious partition problems in graphs of girth at least 4. For the bipartite cut problem we show that every graph G with m edges, whose shortest cycle has length at least $r \geqslant 4$, has a bipartite subgraph with at least $\frac{m}{2} + c(r)m^{\frac{r}{r+1}}$ edges. The order of the error term in this result is shown to be optimal for $r = 5$ thus settling a special case of a conjecture of Erdős. (The result and its optimality for another special case, $r = 4$, were already known.) For judicious partitions, we prove a general result as follows: if a graph $G = (V, E)$ with m edges has a bipartite cut of size $\frac{m}{2} + \delta$, then there exists a partition $V = V_1 \cup V_2$ such that both parts V_1, V_2 span at most $\frac{m}{4} - (1 - o(1))\frac{\delta}{2} + O(\sqrt{m})$ edges for the case $\delta = o(m)$, and at most $(\frac{1}{4} - \Omega(1))m$ edges for $\delta = \Omega(m)$. This enables one to extend results for the bipartite cut problem to the corresponding ones for judicious partitioning.

[1] Research supported in part by a State of New Jersey Grant, by a USA–Israel BSF grant, by a grant from the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.
[2] Research supported in part by NSF Grant DSM 9971788 and DARPA Grant F33615-01-C-1900.
[3] Research supported in part by a USA–Israel BSF Grant, by a grant from the Israel Science Foundation and by a Bergmann Memorial Grant.
[4] Research supported in part by NSF Grants DMS-0106589, CCR-9987845 and by the State of New Jersey.

0095-8956/03/$ - see front matter
doi:10.1016/S0095-8956(03)00036-4

1. Introduction

Many problems in Extremal Graph Theory are instances of the following general setting: given a fixed graph H or a family of fixed graphs $\mathscr{H} = \{H_1, \ldots, H_k\}$ and a large graph $G = (V, E)$ on $|V| = n$ vertices, estimate the extremal values of various graph theoretic parameters of G as functions of n, assuming G is H-free or more generally $(H_1, \ldots, H_k)$-free. Central questions such as those of studying the Turán number $ex(n, H)$ or the Ramsey number $R(H, K^n)$ fall into this category.

In some extremal problems, the size of the large graph $G = (V, E)$ is naturally measured by its number of edges $m = |E|$ rather than by its number of vertices $n = |V|$. Two such problems are the maximal bipartite cut (or Max-Cut) problem, where one seeks to partition the vertex set V into two disjoint parts V_1 and V_2 so that the number of edges of G crossing between V_1 and V_2 is maximal, and the so-called judicious partition problem, where the task is to find a partition $V = V_1 \cup V_2$ such that both parts V_1 and V_2 span the smallest possible number of edges. Formally, for a graph $G = (V, E)$ we define

$$f(G) = \max\{e(V_1, V_2):\ V = V_1 \cup V_2, V_1 \cap V_2 = \emptyset\},$$

$$g(G) = \min_{V = V_1 \cup V_2} \max\{e(V_1), e(V_2)\},$$

where, as usual, $e(U, W)$ is the number of edges of G between the (disjoint) subsets $U, W \subset V$, and $e(U)$ is the number of edges of G spanned by U. Thus, the bipartite cut problem is that of computing the value of $f(G)$, and the judicious partition problem asks to compute $g(G)$. The above two functions are closely connected; moreover, bounding $g(G)$ from above supplies immediately a lower bound for $f(G)$: $f(G) \geqslant m - 2g(G)$. We provide more extensive background information about both these problems later in the paper.

Consider a random partition $V = V_1 \cup V_2$, obtained by assigning each vertex $v \in V$ to V_1 or to V_2 with probability $\frac{1}{2}$ independently. It is easy to see that each edge of G has probability $\frac{1}{2}$ to cross between V_1 and V_2, probability $\frac{1}{4}$ to fall inside V_1, and the same probability $\frac{1}{4}$ to fall inside V_2. It follows that the expected number of edges in the cut (V_1, V_2) is $m/2$, and the expected number of edges in each part V_i is $m/4$. While for the bipartite cut problem the above simple argument shows that every graph G with m edges has a cut of size at least $m/2$, implying $f(G) \geqslant m/2$, for the judicious partitioning it is insufficient to derive $g(G) \leqslant m/4$. Still, it indicates that the right answer should be about $m/2$ for the bipartite cut problem, and about $m/4$ for the judicious partition problem. Therefore, in many cases it is the error term after $m/2$ or $m/4$, respectively, we will be interested in.

In this paper we consider the above two extremal problems when the forbidden graphs H_i are short cycles, or in other words, the graph G is assumed to have girth bounded from below by a parameter r. (Given a graph G, the *girth* of G is the length of the shortest cycle in G; in case G is a forest we set $girth(G) = \infty$). We prove the following results about the bipartite cut problem.

Theorem 1.1. *Let $r \geqslant 4$ be a fixed integer. Then there exists a constant $c > 0$ such that every graph G with m edges and girth at least r satisfies*

$$f(G) \geqslant \frac{m}{2} + cm^{\frac{r}{r+1}}.$$

Theorem 1.2. *There exists an absolute constant $c' > 0$ such that for infinitely many m there exists a graph G with m edges and girth at least 5 for which*

$$f(G) \leqslant \frac{m}{2} + c'm^{\frac{5}{6}}.$$

Thus, the estimate on the error term of Theorem 1.1 is tight up to a constant factor for the case $r = 5$. This settles (in a strong form) a special case of a conjecture of Erdős discussed in more detail in the next section. The assertion of Theorem 1.1 for $r = 4$ and its tightness in this case have been established by the first author in [2].

As for judicious partitions, we prove a very general result, connecting the size of an optimal bipartite cut with the best value of a judicious partition.

Theorem 1.3. *Let $G = (V, E)$ be a graph with m edges whose maximal bipartite cut has cardinality $f(G) = \frac{m}{2} + \delta$. If $\delta \leqslant m/30$, then there exists a partition $V = V_1 \cup V_2$ of the vertex set of G such that*

$$e(V_i) \leqslant \frac{m}{4} - \frac{\delta}{2} + \frac{10\delta^2}{m} + 3\sqrt{m}, \quad i = 1, 2.$$

Therefore, if $\delta = o(m)$ but $\delta \gg \sqrt{m}$, it follows that $g(G) = m/4 - (1 - o(1))\delta/2$. The case of large δ is covered by the following complementary theorem.

Theorem 1.4. *Let $G = (V, E)$ be a graph with m edges whose maximal bipartite cut has cardinality $f(G) = \frac{m}{2} + \delta$. If $\delta \geqslant m/30$ and m is large enough, then there exists a partition $V = V_1 \cup V_2$ of the vertex set of G such that*

$$e(V_i) \leqslant \frac{m}{4} - \frac{m}{100}, \quad i = 1, 2.$$

Combining the above two theorems with Theorem 1.1 we immediately get the following estimate on the judicious partition problem for graphs with given girth:

Corollary 1.5. *Let $r \geqslant 4$ be a fixed integer. Then there exists a constant $c > 0$ such that every graph G with m edges and girth at least r satisfies*

$$g(G) \leqslant \frac{m}{4} - cm^{\frac{r}{r+1}}.$$

Obviously, the above-mentioned tightness results for Theorem 1.1 for $r = 4, 5$ carry over to tightness results for Corollary 1.5.

The rest of the paper is organized as follows. In Section 2, we discuss the bipartite cut problem, first surveying necessary background and then proving Theorems 1.1 and 1.2. Section 3 is devoted to the judicious partition problem. There we first cover relevant previous developments and then prove Theorems 1.3 and 1.4. Section 4, the last section of the paper, contains some concluding remarks and a discussion of related open problems.

In the course of the paper, we will make no serious attempt to optimize the absolute constants involved. For the sake of simplicity of presentation we will drop occasionally floor and ceiling signs whenever these are not crucial.

2. Bipartite cuts

2.1. Background

As we indicated in the introduction, it is quite easy to show that every graph $G = (V, E)$ with m edges contains a bipartite cut (V_1, V_2) spanning at least $m/2$ edges. This elementary result can be improved by providing a more accurate estimate for the error term after the main term $m/2$. Edwards [10,11] proved the essentially best possible result that every graph G with m edges satisfies

$$f(G) \geqslant \frac{m}{2} + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8}.$$

This result is easily seen to be tight in case G is a complete graph on an odd number of vertices, that is, whenever $m = \binom{k}{2}$ for some odd integer k. Estimates on the second error term for other values of m can be found in [2,3,8].

The problem of estimating the minimum possible size of the maximum cut in graphs without short cycles has been raised by Paul Erdős in one of his numerous problem papers [12]. There he introduced the function

$$f_r(m) = \min\{f(G) \colon |E(G)| = m, girth(G) \geqslant r\}$$

and conjectured that for every $r \geqslant 4$ there exists a constant $c_r > 0$ such that for every $\varepsilon > 0$

$$\frac{m}{2} + m^{c_r - \varepsilon} < f_r(m) < \frac{m}{2} + m^{c_r + \varepsilon}$$

provided $m > m(\varepsilon)$. He also mentioned that together with Lovász they proved that

$$\frac{m}{2} + c_2 m^{c_r''} < f_r(m) < \frac{m}{2} + c_1 m^{c_r'},$$

where c_r' and c_r'' are greater than $\frac{1}{2}$ and less than one for all $r > 3$ and tend to one as r tends to infinity. (In this statement, we have corrected an apparent typo in Erdős' paper.)

The case $r = 4$, i.e., the case of triangle-free graphs has attracted most of the attention so far. After a series of papers by various researchers [12,14,16] the first author proved in [2] that if G is a triangle-free graph with m edges, then

$$f(G) \geqslant \frac{m}{2} + cm^{4/5}$$

for some absolute positive constant c. In the same paper [2], the error term of the above estimate is shown to be tight by showing that for every $m > 0$ there exists a triangle-free graph G with m edges for which $f(G) \leqslant \frac{m}{2} + c_0 m^{4/5}$, for an absolute constant $c_0 > 0$. This upper bound is based on a construction of regular triangle-free graphs with extremal spectral properties, given in [1].

Here we generalize the above-stated bounds for the case of graphs of higher girth. The proof of the lower bound of Theorem 1.1, given in the next subsection, utilizes techniques from several previous papers on the subject. We are able to provide a matching upper bound for the case of $r = 5$, i.e., for graphs without 3- and 4-cycles, thus settling the above-mentioned problem of Erdős for this case as well. This result (Theorem 1.2) is proven in Subsection 2.3, where, following the method in [2], we use spectral properties to estimate from above the size of a maximal bipartite cut.

2.2. Lower bound

In this subsection, we obtain a lower bound on the size of the maximum bipartite subgraphs of graphs with girth at least r. We need the following simple lemma from [12], whose short proof is included here for the sake of completeness.

Lemma 2.1. *Let G be a graph with m edges and chromatic number t. Then G contains a bipartite subgraph with at least $\frac{t+1}{2t} m = \frac{m}{2} + \frac{m}{2t}$ edges.*

Proof. Since the chromatic number of G is t we can decompose its vertex set into t independent subsets $V_1, \dots, V_t$. Partition these subsets randomly into two parts, containing $\lfloor \frac{t}{2} \rfloor$ and $\lceil \frac{t}{2} \rceil$ sets V_i, respectively. Let H be a bipartite subgraph of G whose color classes are the above two parts. Note that for every fixed edge e of G the probability that its ends lie in distinct classes of H is

$$\mathbf{Pr}(e \in E(H)) = \frac{\lfloor \frac{t}{2} \rfloor \lceil \frac{t}{2} \rceil}{\binom{t}{2}} \geqslant \frac{\frac{t^2-1}{4}}{\frac{t(t-1)}{2}} = \frac{t+1}{2t}.$$

By linearity of expectation, the expected number of edges in H is at least $\frac{t+1}{2t} m$. This completes the proof. $\square$

Next we need a result of Shearer [16], which provides a very useful lower bound on the size of a maximum bipartite subgraph in a triangle-free graph.

Proposition 2.2. *Let G be a triangle-free graph with m edges, and let $d_1, \ldots, d_n$ be the degrees of the vertices in G. Then*

$$f(G) \geqslant \frac{m}{2} + \frac{1}{8\sqrt{2}} \sum_{i=1}^{n} \sqrt{d_i}.$$

Finally, we shall also use the following upper bound, proved by Bondy and Simonovits [9], on the maximum number of edges in graphs without cycles of a given even length. (We note that in fact we need here only the simpler, similar estimate, for the maximum number of edges in graphs with no short cycles at all, but we include this result as it may be helpful in dealing with the related problem of estimating the maximum cut in graphs without a cycle of a fixed, given length.)

Proposition 2.3. *Let $l \geqslant 2$ be an integer and let G be a graph of order n. If G contains no cycle of length $2l$, then the number of edges in G is at most $100ln^{1+1/l}$.*

Having finished all the necessary preparations we are ready to prove our first theorem.

Proof of Theorem 1.1. To prove the theorem we use the argument from [2] with some additional ideas. We will assume throughout the proof that m is sufficiently large. Let $r \geqslant 4$ be a fixed integer and let G be a graph with n vertices, m edges and with girth at least r. Define $d = \lfloor 100rm^{\frac{2}{r+1}} \rfloor$. First, we consider the case when G has no subgraph with minimum degree greater than d.

In this case, it is easy to see that there exists a labeling $v_1, \ldots, v_n$ of the vertices of G so that for every i, the number of neighbors v_j of v_i with $j < i$ is at most d. Indeed, let v_n be the vertex of minimal degree in G. Clearly, the degree of v_n is at most d, delete it from G and repeat this procedure. Let d_i denote the degree of v_i in G and let d_i' be the number of neighbors v_j of v_i with $j < i$. Obviously, $\sum_{i=1}^{n} d_i' = m$. Since G is triangle-free, by Proposition 2.2 we obtain

$$\begin{aligned} f(G) &\geqslant \frac{m}{2} + \frac{1}{8\sqrt{2}} \sum_{i=1}^{n} \sqrt{d_i} \geqslant \frac{m}{2} + \frac{1}{8\sqrt{2}} \sum_{i=1}^{n} \sqrt{d_i'} \\ &\geqslant \frac{m}{2} + \frac{1}{8\sqrt{2}} \frac{\sum_{i=1}^{n} d_i'}{\sqrt{d}} = \frac{m}{2} + \frac{1}{8\sqrt{2}} \frac{m}{\sqrt{d}} = \frac{m}{2} + \Omega(m^{\frac{r}{r+1}}), \end{aligned}$$

as needed.

Now suppose that there exists a subset of vertices U of G of order u such that the induced subgraph $G[U]$ of G has minimum degree greater than d. We first prove that in this case r should be even. Suppose not, i.e., $r = 2l + 1$ for some integer $l \geqslant 2$. Note that the number of edges in $G[U]$ is at least $ud/2$ and at most the number of edges in G, which is m. This implies that $u \leqslant 2m/d$. In addition, we have that $G[U]$ contains no cycle of length $2l$. Then using the fact that $d = \lfloor 100(2l+1)m^{\frac{1}{l+1}} \rfloor$ together with Proposition 2.2, we conclude that the number of edges in this graph is at most

$$100lu^{1+1/l} \leqslant 100lu\left(\frac{2m}{d}\right)^{1/l} \leqslant 100lu(m^{\frac{l}{l+1}})^{1/l} < \frac{ud}{2},$$

a contradiction. Therefore, in the rest of the proof we will assume that r is even and set $r = 2q + 2$ for some integer $q \geqslant 1$.

Next we prove that U contains a subset U' such that the induced subgraph $G[U']$ spans at least $ud/4$ edges and is t-colorable for $t = \lceil \frac{2u}{d^q} \rceil$. Indeed, let T be a random subset of U obtained by picking uniformly at random, with repetitions allowed, t vertices from U. Let x be a fixed vertex of U. Denote by $S(x)$ the set of vertices in U which are at distance exactly q from x and denote by s_x the size of $S(x)$. Since the minimal degree of $G[U]$ is greater than d and $G[U]$ contains no cycle of length at most $2q + 1$, it is easy to see that $s_x > d^q$ for every $x \in U$. This, together with the definition of t, implies that the probability that $S(x) \cap T$ is empty is at most

$$\left(1 - \frac{s_x}{u}\right)^t < \left(1 - \frac{d^q}{u}\right)^t \leqslant e^{-td^q/u} = e^{-2} < \frac{1}{4}.$$

It follows that for every fixed edge (x, y) of $G[U]$, the probability that both $S(x)$ and $S(y)$ have non-empty intersection with T is at least $\frac{1}{2}$. Let U' be the set of all vertices x in U such that $S(x) \cap T \neq \emptyset$ and let $G[U']$ be the graph induced by U'. By linearity of expectation, the expected number of edges in $G[U']$ is at least $e(U)/2 \geqslant ud/4$. Hence, there exists a particular set T of size at most t such that the corresponding graph $G[U']$ spans at least $e(U') \geqslant ud/4$ edges.

Fix such sets T and U' and define a coloring of $G[U']$ in t colors by coloring each vertex $x \in U'$ by the smallest index of a vertex from T which belongs to $S(x)$. Since $G[U]$ has no cycles of length at most $2q + 1$, it clearly follows that no edge can have both its endpoints at distance exactly q from the same vertex in T. This proves that the coloring defined above is a proper coloring and the set U' with the required properties indeed exists.

Now by Lemma 2.1, there exists a partition of U' into two disjoint subsets U_1 and U_2 so that

$$\begin{aligned} e(U_1, U_2) &\geqslant \frac{e(U')}{2} + \frac{e(U')}{2t} \geqslant \frac{e(U')}{2} + \frac{ud}{8}\left\lceil \frac{2u}{d^q} \right\rceil^{-1} = \frac{e(U')}{2} + \Omega(d^{q+1}) \\ &= \frac{e(U')}{2} + \Omega(d^{r/2}) = \frac{e(U')}{2} + \Omega(m^{\frac{r}{r+1}}). \end{aligned}$$

Now we can assign the remaining vertices in $V(G) - U'$ one by one either to U_1 or to U_2, each time adding a vertex to the subset in which it has more neighbors and breaking ties arbitrarily. This ensures that at least half of the edges which are not in $G[U']$ will lie in the bipartite graph which we obtain in the end of this process. Therefore,

$$f(G) \geqslant \frac{e(G) - e(U')}{2} + \frac{e(U')}{2} + \Omega(m^{\frac{r}{r+1}}) = \frac{m}{2} + \Omega(m^{\frac{r}{r+1}}),$$

completing the proof of the theorem. $\square$

2.3. Graphs with girth 5

In this subsection, we show that the lower bound of Theorem 1.1 is tight, up to a constant factor, for graphs with girth at least 5. To do so we will need the following folklore result, which provides an upper bound for $f(G)$, for a regular graph G, in terms of the smallest eigenvalue of its adjacency matrix. For completeness, we include the short proof.

Lemma 2.4. *Let G be a d-regular graph of order n (which may have loops each of which contributes 1 to the degree of its vertex). Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ be the eigenvalues of the adjacency matrix of G. Then*

$$f(G) \leqslant \frac{dn}{4} - \frac{\lambda_n n}{4}.$$

Proof. Let $V = \{1, \ldots, n\}$ and let $A = (a_{ij})$ be the adjacency matrix of $G = (V, E)$, where a_{ii} corresponds to the number of loops at vertex i. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be any vector with coordinates ± 1. Since the graph G is d-regular we have that $\sum_i a_{ij} = \sum_j a_{ij} = d$ and therefore

$$\sum_{(i,j) \in E} (x_i - x_j)^2 = d \sum_{i=1}^{n} x_i^2 - \sum_{i,j} a_{ij} x_i x_j = dn - \mathbf{x}^t A \mathbf{x}.$$

By the variational definition of the eigenvalues of A, for any vector $z \in R^n$, $z^t A z \geqslant \lambda_n ||z||^2$. Thus,

$$\sum_{(i,j) \in E} (x_i - x_j)^2 = dn - \mathbf{x}^t A \mathbf{x} \leqslant dn - \lambda_n ||\mathbf{x}||^2 = dn - \lambda_n n. \tag{1}$$

Let $V = V_1 \cup V_2$ be an arbitrary partition of V into two disjoint subsets and let $e(V_1, V_2)$ be the number of edges in the bipartite subgraph of G with bipartition (V_1, V_2). For every vertex $v \in V(G)$ define $x_v = 1$ if $v \in V_1$ and $x_v = -1$ if $v \in V_2$. Note that for every edge (i,j) of G, $(x_i - x_j)^2 = 4$ if this edge has its ends in distinct parts of the above partition and is zero otherwise. Now using (1), we conclude that

$$e(V_1, V_2) = \frac{1}{4} \sum_{(i,j) \in E} (x_i - x_j)^2 \leqslant \frac{1}{4} (dn - \lambda_n n) = \frac{dn}{4} - \frac{\lambda_n n}{4}. \qquad \square$$

In order to prove Theorem 1.2 we will use the so-called Erdős–Rényi graph [13], arising from the projective plane $PG_2(p)$ over a finite field. Let p be a prime power and let $\mathbf{F}_p$ be the finite field with p elements. Consider the three-dimensional vector space $\mathbf{F}_p^3$. Two vectors $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in this space are called *orthogonal* if $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3 = 0$, in which case we write $x \perp y$. Similarly, for any two subsets X, Y of $\mathbf{F}_p^3$ we write $X \perp Y$ iff $\langle x, y \rangle = 0$ for any two vectors $x \in X$ and $y \in Y$. Let G be a graph whose vertices are all one-dimensional subspaces of $\mathbf{F}_p^3$. Clearly, the number of vertices of G is $n = (p^3 - 1)/(p - 1) = p^2 + p + 1$ and we denote them by $\mathbf{v}_i$, $1 \leqslant i \leqslant p^2 + p + 1$. Two vertices $\mathbf{v}_i$ and $\mathbf{v}_j$ are adjacent in G if $\mathbf{v}_i \perp \mathbf{v}_j$. Note that G has some vertices with loops and it is easy to see that all its vertices have degree $d = p + 1$. Thus, the sum of the degrees of the vertices in G is $dn = (p + 1)(p^2 + p + 1) = (1 + o(1))n^{3/2}$. Next, we briefly summarize the properties of G we will need later in our proof. This is done in the following simple lemma (which is essentially known).

Lemma 2.5. *Let G be the graph defined above. Then it has the following properties*:

(i) *For every pair of vertices in G there is exactly one vertex of G adjacent to both of them.*
(ii) *The largest eigenvalue of the adjacency matrix of G is $p + 1$ and all other eigenvalues are $\pm\sqrt{p}$.*
(iii) *The set V_0 of all vertices of G with loops has size at most $2(p + 1)$.*

Proof. (i) Let $\mathbf{v}_i, \mathbf{v}_j$ be two distinct vertices of G, then they span a two-dimensional subspace of $\mathbf{F}_p^3$. Thus, the set of vectors orthogonal to $\mathbf{v}_i$ and $\mathbf{v}_j$ has dimension one and corresponds to a unique vertex of G adjacent to both $\mathbf{v}_i$ and $\mathbf{v}_j$.

(ii) Let $A_G = (a_{ij})$ be the adjacency matrix of G, where a_{ii} corresponds to the number of loops at vertex i and let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ be its eigenvalues. Since the graph G is $(p + 1)$-regular we have that $\lambda_1 = p + 1$. Consider now the matrix A_G^2. Clearly, this matrix has eigenvalues $\lambda_1^2, \ldots, \lambda_n^2$. By definition, every vertex of G has at most one loop. Therefore, the diagonal entries of A_G^2 are just the degrees of vertices of G and thus are equal to $p + 1$. In addition, for any $i \neq j$ the ijth entry of this matrix is simply the number of vertices adjacent to both $\mathbf{v}_i$ and $\mathbf{v}_j$ and by (i) is equal to 1. Using this it is easy to deduce that the eigenvalues of A_G^2 are $(p + 1)^2$ with multiplicity one and p with multiplicity $n - 1$. This implies that all eigenvalues of A_G except the first one are $\pm\sqrt{p}$.

(iii) By definition, the size of V_0 is the number of one-dimensional subspaces of $\mathbf{F}_p^3$ which are self-orthogonal. Note that any vector (x, y, z) in $\mathbf{F}_p^3$, which is self-orthogonal satisfies the equation $x^2 + y^2 + z^2 = 0$ over $\mathbf{F}_p$. Since for every choice of x and y we can have at most two values for z which will satisfy the equation, we obtain that the number of non-zero solutions of this equation is at most $2(p^2 - 1)$. Since every one-dimensional self-orthogonal subspace contains $p - 1$ such solutions and no solution is contained in two different subspaces we conclude that $|V_0| \leqslant \frac{2(p^2-1)}{p-1} = 2(p + 1)$. This completes the proof. □

Remark. Actually, one can show that $|V_0| = p + 1$, but for our purposes it is enough to have the above weaker bound which is easier to prove.

Let G be the graph constructed above. From assertion (i) of Lemma 2.5 it follows immediately that G contains no cycles of length 4. In addition, every edge $(\mathbf{v}_i, \mathbf{v}_j)$ of this graph, for which $\mathbf{v}_i, \mathbf{v}_j \notin V_0$, is contained in some cycle of length 3. Indeed, in this case $\mathbf{v}_i, \mathbf{v}_j$ have a common neighbor which is distinct from both of them. Also, using Lemma 2.4 we have

$$f(G) \leqslant \frac{dn}{4} - \frac{\lambda_n n}{4} \leqslant \frac{dn}{4} + \frac{\sqrt{p}n}{4} = \frac{dn}{4} + O(n^{5/4}).$$

Let H be the graph obtained from G by deleting all edges of G adjacent to vertices in V_0, i.e., edges not contained in any cycle of length 3. By definition, H is a graph of order n which has at least

$$e(H) \geqslant \frac{dn - 2(p+1)|V_0|}{2} \geqslant \frac{dn}{2} - 2(p+1)^2 = \frac{dn}{2} - O(n) = (1/2 + o(1))n^{3/2}$$

edges. Every edge of H is contained in some cycle of length 3 and the maximum bipartite subgraph of H still has size at most

$$f(H) \leqslant f(G) \leqslant \frac{dn}{4} + O(n^{5/4}) = \frac{e(H)}{2} + O(n) + O(n^{5/4}) = \frac{e(H)}{2} + O(n^{5/4}).$$

Hence, to complete the proof of Theorem 1.2 we need to prove the lemma below.

Lemma 2.6. *Let H be a graph of order n with $e = (1/2 + o(1))n^{3/2}$ edges and with the following properties*:

- *H has no cycles of length* 4;
- *every edge of H is contained in some triangle*, *i.e.*, *cycle of length* 3;
- $f(H) \leqslant \frac{e}{2} + O(n^{5/4})$.

Then H contains a subgraph H_0 with $m = 2e/3 = (1/3 + o(1))n^{3/2}$ edges and girth at least 5, *for which*

$$f(H_0) \leqslant \frac{m}{2} + O(n^{5/4}) = \frac{m}{2} + O(m^{5/6}).$$

Proof. First note that since H has no cycle of length 4 every two triangles in H are edge disjoint. Since every edge of this graph is contained in some triangle we conclude that the set of edges of H is a union of $e/3$ edge disjoint triangles. Let H_0 be a subgraph of H obtained by deleting uniformly at random one edge from every triangle in H. Clearly the number of edges in H_0 is $2e/3$, since H_0 contains precisely two edges from every triangle in H. In addition, H_0 is triangle-free, since we destroyed all triangles in H. This implies that the girth of H_0 is at least 5.

Next we show that with probability $1 - o(1)$, the new graph contains no large bipartite subgraphs and thus satisfies the assertion of the lemma. Indeed, let $V(H) = V_1 \cup V_2$ be an arbitrary partition of V into two disjoint subsets and let $t = e_H(V_1, V_2)$ be the number of edges in the corresponding bipartite subgraph of H. Note that for every triangle in H either none or two of its edges belong to the cut

(V_1, V_2). It follows that we can find a set $C_1, \ldots, C_{t/2}$ of edge disjoint triangles such that every C_i contains precisely two edges from the cut (V_1, V_2). Recall that for every triangle C_i, $1 \leqslant i \leqslant t/2$ we deleted one of its edges uniformly at random. Let x_i', $1 \leqslant i \leqslant t/2$, be the random variable equal to the number of edges of the triangle C_i that belong to the cut (V_1, V_2) and were not deleted and let $x_i = x_i' - 1$. By definition, we have that $x_i = 1$ with probability $\frac{1}{3}$ (i.e., in case when we delete the edge of C_i not in the cut) and $x_i = 0$ with probability $\frac{2}{3}$ (i.e., in case when we delete one of the two edges of C_i that are in the cut). Clearly, the total number of edges of the graph H_0 in the cut (V_1, V_2) equals $e_{H_0}(V_1, V_2) = \sum_{i=1}^{t/2} x_i' = t/2 + \sum_{i=1}^{t/2} x_i$. Since $X = \sum_{i=1}^{t/2} x_i = e_{H_0}(V_1, V_2) - t/2$ is a binomially distributed random variable with parameters $t/2$ and $\frac{1}{3}$, it follows by the standard estimates for Binomial distributions (see, e.g., [5, Appendix A]) that

$$\mathbf{Pr}\left(X - \frac{t}{6} > a = cn^{5/4}\right) \leqslant e^{-\Omega(a^2/t)} = e^{-\Omega(c^2 n^{5/2}/t)}.$$

Choosing c large enough and using the fact that $t \leqslant m \leqslant O(n^{3/2})$ we conclude that

$$\mathbf{Pr}\left(e_{H_0}(V_1, V_2) - \frac{2}{3}t > cn^{5/4}\right) = \mathbf{Pr}\left(X - \frac{t}{6} > cn^{5/4}\right) < e^{-n}.$$

Since the total number of partitions of H is at most 2^n, this implies that with probability $1 - o(1)$ for every partition $V = V_1 \cup V_2$ we have

$$e_{H_0}(V_1, V_2) \leqslant \tfrac{2}{3} e_H(V_1, V_2) + O(n^{5/4}).$$

In particular, since the number of edges in H_0 is $m = 2e/3 = (\frac{1}{3} + o(1))n^{3/2}$ we obtain that with probability $1 - o(1)$ the size of a maximum bipartite subgraph of H_0 satisfies

$$\begin{aligned} f(H_0) &\leqslant \frac{2}{3} f(H) + O(n^{5/4}) = \frac{2}{3}\left(\frac{e}{2} + O(n^{5/4})\right) + O(n^{5/4}) \\ &= \frac{m}{2} + O(n^{5/4}) = \frac{m}{2} + O(m^{5/6}). \end{aligned}$$

This completes the proof of the lemma. □

In fact, relying on known results on distances between consecutive primes (see, e.g., [6]), one may prove that the assertion of Theorem 1.2 holds for *all* m. To show this, we can take, for a given m, several disjoint copies (of varying sizes) of the graph $H_0 = H_0(p)$ constructed in Lemma 2.6 so that their total number of edges is less than m and is at least $m - o(m^{5/6})$, and then add, if necessary, some isolated edges to create a graph G with girth at least 5 and m edges, satisfying

$$f(G) \leqslant \frac{m}{2} + c' m^{\frac{5}{6}}.$$

This shows that for $r = 5$ the exponent $\frac{5}{6}$ in Theorem 1.1 cannot be improved.

3. Judicious partitions

3.1. Background

It is easy to prove that a partition (V_1, V_2) of a graph $G = (V, E)$ with m edges, for which every vertex $v \in V_i$ has at least as many neighbors of the opposite part V_{3-i} as of its own part, is such that $e(V_1), e(V_2) \leqslant \frac{1}{2} e(V_1, V_2)$, and therefore $e(V_1), e(V_2) \leqslant m/3$. Since a partition with the maximal number of crossing edges clearly has the above property, we get that $g(G) \leqslant m/3$. This bound is optimal as shown by the example of a complete graph K_3. However, for large values of m one can expect to do much better. The probabilistic reasoning, described in the introduction, indicates that the right answer for growing m should be around $m/4$. Indeed, Porter [15] proved in 1992 that every graph with $m \geqslant 1$ edges has a bipartition in which each class contains at most $m/4 + \sqrt{m/8}$ edges. The best possible bound for a general graph has been obtained by the second author and Scott in [7], where it was proved that for a graph G with m edges,

$$g(G) \leqslant \frac{m}{4} + \sqrt{\frac{m}{32} + \frac{1}{256}} - \frac{1}{16},$$

i.e., exactly one half of the Edwards bound for bipartite cuts. (In fact, it was proven in [7] that there exists a partition (V_1, V_2) meeting both the bound of Edwards for bipartite cuts and the above stated bound for judicious partitions). This bound is exact for complete graphs of odd order. To the best of our knowledge, the judicious partitioning problem for graphs with forbidden subgraphs has not been considered in the literature.

The problems of bounding bipartite cuts and judicious partitions are closely related. Hence, a rather natural approach to the (probably more complicated) judicious partitioning problem would be to derive bounds for judicious partitions from those on bipartite cuts. This approach is carried out in our Theorem 1.3, where it is proven that if a general graph G with m edges has a bipartite cut with $m/2 + \delta$ edges, i.e., with a surplus $\delta = o(m)$ over the trivial $m/2$ bound, then this surplus can be divided almost equally between the two parts of the cut, resulting in a partition in which both parts span about $m/4 - \delta/2 + o(\delta) + O(\sqrt{m})$ edges. (Observe that the $O(\sqrt{m})$ correction term is needed in this estimate due to the optimality of the above-stated result of [7]). Moreover, as we are about to show, the proof starts with an optimal bipartite cut and proceeds by moving vertices between the two parts V_1 and V_2 so as to balance the number of edges spanned by them, while maintaining the almost optimality of the bipartite cut between V_1 and V_2. For the case of δ linear in m, Theorem 1.4 shows that $g(G)$ is smaller than $m/4$ by an additive factor linear in m. Thus, Theorems 1.3 and 1.4 form a bridge between the two problems considered in this paper and enable one to derive results on the judicious partition problem by looking at the corresponding bipartite cut problem. Combining this with Theorem 1.1 results in Corollary 1.5, bounding from above the value of an optimal judicious partition in graphs without short cycles.

The proofs of Theorems 1.3 and 1.4 are given in the next subsection.

3.2. Proofs of Theorems 1.3 and 1.4

For a vertex $v \in V$ and a subset $U \subseteq V$ we denote by $d(v, U)$ the number of neighbors of v in U.

Proof of Theorem 1.3. The main ingredient of the proof is the following lemma.

Lemma 3.1. *Let $G = (V, E)$ be a graph with m edges and with $f(G) = \frac{m}{2} + \delta$, where $\delta \leqslant \frac{m}{30}$. Suppose $V = V_1 \cup V_2$ is a partition of $V(G)$ for which $d(v, V_1) \leqslant d(v, V_2)$ for every vertex $v \in V_1$. If $e(V_1) \geqslant \frac{m}{4} - \frac{\delta}{2}$, then there exists a vertex $v \in V_1$ such that $d(v, V_1) \leqslant 3\sqrt{m}$ and $d(v, V_2) \leqslant (1 + \frac{10\delta}{m}) d(v, V_1)$.*

Proof. We prove the lemma by showing that the total degree of vertices of V_1 violating any of the required conditions does not reach the total degree of vertices in V_1.

Define $T_1 = \{v \in V_1 \colon d(v, V_1) > 3\sqrt{m}\}$. Observe that as $d(v, V_1) \leqslant d(v, V_2)$ for every vertex $v \in V_1$, it follows that

$$2e(V_1) = \sum_{v \in V_1} d(v, V_1) \leqslant \sum_{v \in V_1} d(v, V_2) = e(V_1, V_2),$$

implying $e(V_1) \leqslant m/3$. Thus, $|T_1| \leqslant 2e(V_1)/(3\sqrt{m}) \leqslant 2\sqrt{m}/9$. Therefore, the set T_1 spans at most $2m/81$ edges. As in the summation $\sum_{v \in T_1} d(v, V_1)$, the edges spanned by T_1 are counted twice and every other edge inside V_1 is counted at most once, we get

$$\sum_{v \in T_1} d(v, V_1) \leqslant e(V_1) + e(T_1) \leqslant e(V_1) + \frac{2m}{81}. \tag{2}$$

Define now $T_2 = \{v \in V_1 \colon d(v, V_2) > (1 + \frac{10\delta}{m}) d(v, V_1)\}$. Then

$$\begin{aligned} e(V_1, V_2) &= \sum_{v \in T_2} d(v, V_2) + \sum_{v \in V_1 \setminus T_2} d(v, V_2) \\ &\geqslant \left(1 + \frac{10\delta}{m}\right) \sum_{v \in T_2} d(v, V_1) + \sum_{v \in V_1 \setminus T_2} d(v, V_1) \\ &= \sum_{v \in V_1} d(v, V_1) + \frac{10\delta}{m} \sum_{v \in T_2} d(v, V_1) = 2e(V_1) + \frac{10\delta}{m} \sum_{v \in T_2} d(v, V_1), \end{aligned}$$

implying

$$\sum_{v \in T_2} d(v, V_1) \leqslant \frac{m}{10\delta} (e(V_1, V_2) - 2e(V_1)).$$

Observe that $e(V_1, V_2) \leqslant f(G) = \frac{m}{2} + \delta$ and that by the lemma assumption $e(V_1) \geqslant \frac{m}{4} - \frac{\delta}{2}$. Hence,

$$\sum_{v \in T_2} d(v, V_1) \leqslant \frac{m}{10\delta}\left(\frac{m}{2} + \delta - 2\left(\frac{m}{4} - \frac{\delta}{2}\right)\right) = \frac{m}{5}. \tag{3}$$

From (2) and (3) we derive

$$\sum_{v \in T_1 \cup T_2} d(v, V_1) \leqslant e(V_1) + \frac{2m}{81} + \frac{m}{5} < e(V_1) + 0.23m. \tag{4}$$

On the other hand, recalling our assumption on δ, we can see that

$$\sum_{v \in V_1} d(v, V_1) = 2e(V_1) \geqslant e(V_1) + \frac{m}{4} - \frac{\delta}{2} \geqslant e(V_1) + \frac{m}{4} - \frac{m}{60} > e(V_1) + 0.23m. \tag{5}$$

Comparing (4) and (5) shows that not all vertices of V_1 are in the union of T_1 and T_2. It follows from the definitions of T_1 and T_2 that any vertex in $V_1 \backslash (T_1 \cup T_2)$ meets the requirements of the lemma. $\square$

We now prove Theorem 1.3. Let $V = U_1 \cup U_2$ be a partition of V satisfying $e(U_1, U_2) = f(G) = \frac{m}{2} + \delta$ and $e(U_1) \geqslant e(U_2)$. Clearly for every vertex $u \in U_1$, $d(u, U_1) \leqslant d(u, U_2)$, as otherwise moving u from U_1 to U_2 would create a bipartite cut of size larger than $e(U_1, U_2) = f(G)$. We will achieve a partition with the desired properties by starting from (U_1, U_2) and by moving a number of vertices from U_1 to U_2 in order to balance the number of edges spanned by those subsets. Lemma 3.1 will help us to maintain the size of the cut almost unchanged. Formally, we start by assigning $V_1 = U_1$, $V_2 = U_2$. Then, as long as $e(V_1) \geqslant \frac{m}{4} - \frac{\delta}{2} + 3\sqrt{m}$, we find a vertex $v_i \in V_1$, for which $d(v_i, V_1) \leqslant 3\sqrt{m}$ and $d(v_i, V_2) \leqslant (1 + \frac{10\delta}{m}) d(v_i, V_1)$ and transfer it to V_2. It is easy to see that the conditions of Lemma 3.1 still apply and therefore such a vertex indeed can be found. We denote $d(v_i, V_1) = a_i$, $d(v_i, V_2) = b_i$. Note that $b_i \leqslant (1 + \frac{10\delta}{m}) a_i$.

Let us look at the final partition (V_1, V_2) after the above-described process has terminated. Suppose the vertices moved from V_1 to V_2 are $v_1, \ldots, v_t$. Clearly,

$$e(V_1) < \frac{m}{4} - \frac{\delta}{2} + 3\sqrt{m}. \tag{6}$$

We now estimate from above the number of edges in V_2. To this end, denote $e(U_1) = m_1$, then $e(U_2) = m - e(U_1, U_2) - e(U_1) = \frac{m}{2} - \delta - m_1$. As $2e(U_1) \leqslant e(U_1, U_2) = \frac{m}{2} + \delta$, we get $m_1 \leqslant \frac{m}{4} + \frac{\delta}{2}$. Notice that while moving a vertex v_i from V_1 to V_2 during the process, we deleted a_i edges from V_1 and added b_i edges to V_2. Therefore, for the final partition (V_1, V_2) we get

$$e(V_1) = e(U_1) - \sum_{i=1}^{t} a_i = m_1 - \sum_{i=1}^{t} a_i, \tag{7}$$

$$e(V_2) = e(U_2) + \sum_{i=1}^{t} b_i = \frac{m}{2} - \delta - m_1 + \sum_{i=1}^{t} b_i \leqslant \frac{m}{2} - \delta - m_1 + \left(1 + \frac{10\delta}{m}\right) \sum_{i=1}^{t} a_i. \tag{8}$$

As each time we moved from V_1 to V_2 a vertex v_i with $d(v_i, V_1) \leqslant 3\sqrt{m}$, it follows that in the final partition (V_1, V_2), $e(V_1) \geqslant \frac{m}{4} - \frac{\delta}{2}$, since (6) was violated just before the last step. Hence, from (7)

$$\sum_{i=1}^{t} a_i = m_1 - e(V_1) \leqslant m_1 - \frac{m}{4} + \frac{\delta}{2}.$$

Therefore, it follows from (8) that

$$\begin{aligned} e(V_2) &\leqslant \frac{m}{2} - \delta - m_1 + \left(1 + \frac{10\delta}{m}\right)\left(m_1 - \frac{m}{4} + \frac{\delta}{2}\right) \\ &= \frac{m}{4} - \frac{\delta}{2} + \frac{10\delta}{m}\left(m_1 - \frac{m}{4} + \frac{\delta}{2}\right) \\ &\leqslant \frac{m}{4} - \frac{\delta}{2} + \frac{10\delta^2}{m}. \end{aligned}$$

This together with (6) establishes the theorem. □

Proof of Theorem 1.4. The proof here is similar to that of Theorem 1.3, with parameters tuned so as to guarantee the error term $m/100$.

We claim that the desired partition can be obtained using the following procedure. Start with an optimal bipartite cut $V = U_1 \cup U_2$, for which $e(U_1, U_2) = f(G) = \frac{m}{2} + \delta$ and $e(U_1) \geqslant e(U_2)$. Initialize $V_1 = U_1$, $V_2 = U_2$, and then, as long as $e(V_1) > m/4 - m/100$ and V_1 contains a vertex v_i for which

$$d(v_i, V_1) \leqslant m/400 \tag{9}$$

and

$$d(v_i, V_2) \leqslant \left(1 + \frac{\delta + \frac{m}{50}}{\frac{23m}{100}}\right) d(v_i, V_1), \tag{10}$$

move v_i to V_2.

Let us show first that the algorithm terminates successfully, i.e., reaches the stage where $e(V_1) \leqslant \frac{m}{4} - \frac{m}{100}$. To do so we need to show that as long as the last condition is not fulfilled a required vertex $v_i \in V_1$, satisfying conditions (9) and (10) exists. Suppose we are at some intermediate stage and the current partition is (V_1, V_2). Define $T_1 = \{v \in V_1 : d(v, V_1) \geqslant m/400\}$. Then, as $e(V_1) \leqslant m/3$, $|T_1| \leqslant 2e(V_1)/(m/400) \leqslant (2m/3)/(m/400) = \frac{800}{3}$, and therefore T spans at most

$(\frac{800}{3})^2/2 < 36\,000$ edges. Hence, similarly to the proof of Theorem 1.3,

$$\sum_{v \in T_1} d(v, V_1) \leqslant e(V_1) + e(T_1) < e(V_1) + 36\,000. \tag{11}$$

Set now

$$T_2 = \left\{ v \in V_1 \colon d(v, V_2) > \left(1 + \frac{\delta + \frac{m}{50}}{\frac{23m}{100}} \right) d(v, V_1) \right\}.$$

Then, again as in the proof of Theorem 1.3, we get

$$\sum_{v \in T_2} d(v, V_1) \leqslant \frac{\frac{23m}{100}}{\delta + \frac{m}{50}} (e(V_1, V_2) - 2e(V_1))$$
$$\leqslant \frac{\frac{23m}{100}}{\delta + \frac{m}{50}} \left(\frac{m}{2} + \delta - 2\left(\frac{m}{4} - \frac{m}{100} \right) \right) = \frac{23m}{100}. \tag{12}$$

Therefore, from (11) and (12) we get

$$\sum_{v \in T_1 \cup T_2} d(v, V_1) < e(V_1) + 36\,000 + \frac{23m}{100} < e(V_1) + 0.24m < 2e(V_1)$$

for sufficiently large m, and hence $V_1 \backslash (T_1 \cup T_2) \neq \emptyset$, implying the existence of a vertex with the required properties.

Let us now estimate the number of edges spanned by the final sets V_1 and V_2. Obviously,

$$e(V_1) \leqslant \frac{m}{4} - \frac{m}{100}. \tag{13}$$

Denote $e(U_1) = m_1$, then $m_1 \leqslant e(U_1, U_2)/2 = \frac{m}{4} + \frac{\delta}{2}$. Suppose we transferred from V_1 to V_2 vertices $v_1, \ldots, v_t$, whose degrees (at the time of movement) were $a_i = d(v_i, V_1)$ and $b_i = d(v_i, V_2)$. As in the end $e(V_1) \geqslant \frac{m}{4} - \frac{m}{100} - \frac{m}{400} = \frac{19m}{80}$, we get

$$\sum_{i=1}^{t} a_i \leqslant m_1 - \frac{19m}{80},$$

implying

$$\sum_{i=1}^{t} b_i \leqslant \left(1 + \frac{\delta + \frac{m}{50}}{\frac{23m}{100}} \right) \left(m_1 - \frac{19m}{80} \right).$$

Therefore,

$$e(V_2) = \frac{m}{2} - \delta - m_1 + \sum_{i=1}^{t} b_i < \frac{m}{2} - \delta - m_1 + \left(1 + \frac{\delta + \frac{m}{50}}{\frac{23m}{100}} \right) \left(m_1 - \frac{19m}{80} \right)$$
$$= \frac{21m}{80} - \delta + \frac{(\delta + \frac{m}{50})(m_1 - \frac{19m}{80})}{\frac{23m}{100}}$$
$$\leqslant \frac{21m}{80} - \delta + \frac{(\delta + \frac{m}{50})(\frac{\delta}{2} + \frac{m}{80})}{\frac{23m}{100}}.$$

We may assume that $\delta \leqslant \frac{13m}{50}$, as otherwise the initial partition (U_1, U_2) satisfies the theorem requirements. An easy check shows that for every δ in the interval $[\frac{m}{30}, \frac{13m}{50}]$ the bound on $e(V_2)$ from the last display, viewed as a quadratic function of the parameter δ, is strictly less than $0.24m$. This together with (13) completes the proof of Theorem 1.4. $\square$

4. Concluding remarks

- The following strengthening of the conjecture of Erdős seems plausible:

  Conjecture 4.1. *For every $r \geqslant 4$, there exist $c_1 = c_1(r), c_2 = c_2(r) > 0$ so that*

  $$\frac{m}{2} + c_1 m^{\frac{r}{r+1}} < f_r(m) < \frac{m}{2} + c_2 m^{\frac{r}{r+1}}.$$

  Note that by Theorem 1.1 the lower bound indeed holds, and by the results of [2] and by our results here, the upper bound also holds for $r = 4, 5$. Moreover, the construction in [1] can be generalized to provide, for every even value of r, graphs with m edges in which the maximum bipartite subgraph is of size at most $\frac{m}{2} + c_3 m^{\frac{r}{r+1}}$, which contain no *odd* cycles of length smaller than r. Unfortunately, these graphs do have short even cycles, and therefore do not prove the upper bound of the above conjecture as stated, though they do provide further indication that its assertion holds.
- It is not difficult to use some of the techniques given here and show that for every fixed graph H there exists a constant $\varepsilon = \varepsilon(H) > 0$ such that for any H-free graph G with m edges $f(G) \geqslant \frac{m}{2} + \Omega(m^{1/2+\varepsilon})$. (One can for example first show that the chromatic number of a K^r-free graph G with m edges satisfies $\chi(G) = O(m^{1/2-\delta})$ for some $\delta = \delta(r) > 0$ by applying known bounds on the off-diagonal Ramsey numbers $R(K^r, K^n)$, and then invoke Lemma 2.1.) Using the results in [4] we can obtain some explicit reasonable estimates for certain specific graphs H. However, we suspect that in fact much more is true, and for any H-free graph G with m edges, $f(G) \geqslant \frac{m}{2} + \Omega(m^{3/4+\varepsilon})$. It is worth noting that the random graph $G = G(n,p)$, satisfies, almost surely, $f(G) \geqslant \frac{n^2 p}{4} + \Omega(n\sqrt{np})$ for every $p = p(n)$ satisfying, say, $p \leqslant \frac{1}{2}$. To see that this is the case fix an ordering $v_1, v_2, \ldots, v_n$ of the set of vertices V of G, and construct the cut $V = V_1 \cup V_2$ greedily, by putting each vertex v_i in its turn in the part which adds more edges to the constructed bipartite graph. Since we can expose the edges from v_i to all previous vertices only after we have already partitioned these vertices, there is an expected discrepancy of $\Omega(\sqrt{(i-1)p})$ between the number of edges from v_i to the two parts constructed so far, implying the desired estimate. Note that even for $p = \frac{1}{2}$ this gives that almost surely $f(G) = \frac{n^2}{4} + \Omega(n^{3/2}) = \frac{m}{2} + \Omega(m^{3/4})$, and it is easy to see that the order of the error term here (and for all other reasonable values of p) is tight.

Acknowledgments

We would like to thank two anonymous referees for many helpful comments.

References

[1] N. Alon, Explicit Ramsey graphs and orthonormal labelings, Electronic J. Combin. 1 (1994) R12, 8pp.
[2] N. Alon, Bipartite subgraphs, Combinatorica 16 (1996) 301–311.
[3] N. Alon, E. Halperin, Bipartite subgraphs of integer weighted graphs, Discrete Math. 181 (1998) 19–29.
[4] N. Alon, M. Krivelevich, B. Sudakov, Turán numbers of bipartite graphs and related Ramsey type questions, submitted for publication.
[5] N. Alon, J. Spencer, The Probabilistic Method, 2nd Edition, Wiley, New York, 2000.
[6] R.C. Baker, G. Harman, J. Pintz, The difference between consecutive primes II, Proc. London Math. Soc. 83 (2001) 532–562.
[7] B. Bollobás, A.D. Scott, Exact bounds for judicious partitions of graphs, Combinatorica 19 (1999) 473–486.
[8] B. Bollobás, A.D. Scott, Better bounds for max cut, in: B. Bollobás (Ed.), Contemporary Combinatorics, Bolyai Society Mathematical Studies, Springer, Berlin, 2002, pp. 185–246.
[9] A. Bondy, M. Simonovits, Cycles of even length in graphs, J. Combin. Theory Ser. B 16 (1974) 97–105.
[10] C.S. Edwards, Some extremal properties of bipartite subgraphs, Canad. J. Math. 3 (1973) 475–485.
[11] C.S. Edwards, An improved lower bound for the number of edges in a largest bipartite subgraph, Proceedings of Second Czechoslovak Symposium on Graph Theory, Prague, 1975, pp. 167–181.
[12] 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, pp. 153–163.
[13] P. Erdős, A. Rényi, On a problem in the theory of graphs (in Hungarian), Publ. Math. Inst. Hungar. Acad. Sci. 7 (1962) 215–235.
[14] S. Poljak, Zs. Tuza, Bipartite subgraphs of triangle-free graphs, SIAM J. Discrete Math. 7 (1994) 307–313.
[15] T.D. Porter, On a bottleneck bipartition conjecture of Erdős, Combinatorica 12 (1992) 317–321.
[16] J. Shearer, A note on bipartite subgraphs of triangle-free graphs, Random Structures Algorithms 3 (1992) 223–226.