Discrete Mathematics 313 (2013) 665–676

Contents lists available at SciVerse ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Variations on a theme of Graham and Pollak

Sebastian M. Cioabă [a,*], Michael Tait [b]

[a] Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA
[b] Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 12 September 2011
Received in revised form 3 December 2012
Accepted 4 December 2012

Keywords:
Biclique
Covering
Hypergraph
Eigenvalue
Rank

ABSTRACT

Graham and Pollak proved that one needs at least $n - 1$ complete bipartite subgraphs (bicliques) to partition the edge set of the complete graph on n vertices. In this paper, we study the generalizations of their result to coverings of graphs with specified multiplicities and to complete uniform hypergraphs. We also discuss the recently disproved Alon–Saks–Seymour Conjecture (which proposed a generalization of the previous result of Graham and Pollak) and compute the exact values of the ranks of the adjacency matrices of the known counterexamples to the Alon–Saks–Seymour Conjecture. The rank of the adjacency matrix of a graph G is related to important problems in computational complexity and provides a non-trivial lower bound for the minimum number of bicliques that partition the edge set of G.

1. Introduction

The *biclique partition number*, $\mathrm{bp}(G)$, of a finite loopless multigraph G is the minimum number of complete bipartite subgraphs (bicliques) whose edges partition the edge set of G. Determining $\mathrm{bp}(G)$ for a general graph G is a hard problem (see Kratzke, Reznick and West [19]). Since the edge set of G can be partitioned into stars centered at the vertices of a vertex cover, $\mathrm{bp}(G)$ is at most the minimum size of a vertex cover of G. Witsenhausen (cf. Graham and Pollak [15,16]) proved that $\mathrm{bp}(G) \geq \max(n_+(G), n_-(G))$, where $n_+(G)$ ($n_-(G)$) is the number of positive (negative) eigenvalues of the adjacency matrix of G. A graph is called *eigensharp* if equality holds in this bound (see [19]).

Graham and Pollak [15,16] studied the biclique partition number of multigraphs in connection with some network routing problems (see also Babai and Frankl [6, Section 1.4], Van Lint [29] and Van Lint and Wilson [30, Chapter 9]). Graham and Pollak observed that the minimum length of $\{0, 1, *\}$-words needed to label the vertices of a connected graph G so that the distance between any two vertices is the number of positions in their labels where one has 1 and the other has 0 equals $\mathrm{bp}(D(G))$. The distance multigraph $D(G)$ has the same vertex set as G, and the multiplicity of the edge uv in $D(G)$ equals the distance between u and v in G. When G is the complete graph K_n on n vertices, $D(K_n) = K_n$, and hence addressing K_n is equivalent to determining $\mathrm{bp}(K_n)$. Using algebraic methods involving Sylvester's law of inertia, Graham and Pollak [16] proved that $\mathrm{bp}(K_n) \geq n-1$. Since the edges of K_n can be partitioned into $n-1$ bicliques (one can use $n-1$ edge disjoint stars, but there are many other ways; see Babai and Frankl [6, Ex 1.4.5, p. 29]), this shows $\mathrm{bp}(K_n) = n-1$. Peck [23], Tverberg [28], and Vishwanathan [32,33] gave other proofs that $\mathrm{bp}(K_n) = n - 1$.

A natural generalization of the Graham–Pollak Theorem asks whether any graph G can be properly colored with $\mathrm{bp}(G)+1$ colors. This result was actually conjectured by Alon, Saks, and Seymour (cf. Kahn [18]). The Alon–Saks–Seymour Conjecture remained open for twenty years until recently when Huang and Sudakov [17] disproved it and constructed graphs G with arbitrarily large biclique partition number such that $\chi(G) \geq c(\mathrm{bp}(G))^{6/5}$, for some absolute positive constant c. In [10], the authors extended Huang and Sudakov's work and constructed other counterexamples to the Alon–Saks–Seymour Conjecture.

* Corresponding author.
E-mail addresses: cioaba@math.udel.edu (S.M. Cioabă), mtait@math.ucsd.edu (M. Tait).

0012-365X/$ – see front matter
doi:10.1016/j.disc.2012.12.001

At the present time, the construction of Huang and Sudakov gives the largest gap between bp(G) and $\chi(G)$, and these authors conjectured in [17] that there exist graphs G with biclique partition number k and chromatic number at least $2^{c\log^2 k}$, for some constant $c > 0$. The existence of such graphs would resolve the complexity of the clique vs. independent set problem (see [17]).

The graphs constructed in [10] are also counterexamples to the Rank-Coloring Conjecture [31] stating that $\chi(G) \leq \text{rank}(A(G))$, where rank$(A(G))$ is the rank of the adjacency matrix, $A(G)$, of G. The Rank-Coloring Conjecture was first disproved by Alon and Seymour [4], and the first superlinear gap between $\chi(G)$ and rank$(A(G))$ was obtained by Razborov [25]. The construction from [10] is also an extension of Razborov's work. At the present time, the biggest gap between $\chi(G)$ and rank$(A(G))$ is given by a construction of Nisan and Wigderson [22]. The Alon–Saks–Seymour Conjecture and the Rank-Coloring Conjecture are closely related to problems in computational complexity (see [17,22,25]).

In this paper, we discuss several extensions of the previous result of Graham and Pollak. In Section 2, we study the minimum number of bicliques needed to cover the edges of a graph G such that the number of times each edge of G is covered belongs to a specific list. If L is a list of positive integers and G is a graph, let $\text{bp}_L(G)$ denote the minimum number of bicliques that partition the edges of G such that each edge of G is contained in exactly l bicliques for some $l \in L$. As each edge is a biclique, this parameter is well-defined, and the Graham–Pollak Theorem can be restated as $\text{bp}_L(K_n) = n - 1$ for $L = \{1\}$. In Section 2, we obtain some lower bounds on $\text{bp}_L(G)$ using algebraic methods (Theorems 2.1 and 2.3), and we present some old and new constructive upper bounds for $\text{bp}_L(K_n)$ for several lists L (Examples 1–6).

In Section 3, we study the hypergraph version of the Graham–Pollak Theorem, where few exact results are known. Let $f_r(n)$ be the minimum number of complete r-partite r-uniform hypergraphs needed to partition the edge set of the complete r-uniform hypergraph on n vertices. The Graham–Pollak Theorem states that $f_2(n) = n-1$. Alon [1] proved that $f_3(n) = n-2$ and $f_r(n) = \Theta(n^{\lfloor r/2 \rfloor})$ for fixed $r \geq 4$. Cioabă, Kündgen, and Verstraëte [9] improved Alon's bounds in the lower order terms. For $r \geq 4$, not many values of $f_r(n)$ are known. In Section 3, we improve the upper bounds due to Alon [1] and to Cioabă, Kündgen, and Verstraëte [9] (Theorem 3.1 and Proposition 3.3), and we determine the exact value of $f_r(r+2)$ for every r (Theorem 3.7) and the asymptotic value of $f_r(r+t)$ for fixed t (Proposition 3.8). The results of this section show that finding the exact value of $f_r(n)$ or even coming up with a conjectured value of $f_r(n)$ will be quite difficult, since $f_r(n)$ does not equal the current lower or upper bounds for $r \geq 6$.

In Section 4, we discuss the recently disproved Alon–Saks–Seymour Conjecture and compute the precise value of the rank of the adjacency matrix of the known counterexamples to this conjecture (Theorem 4.1 and Proposition 4.2). Our results extend the work from [10] and give the exact order of magnitude of the biclique partition number of these counterexamples. We conclude the paper with some open problems.

2. *L*-bipartite coverings

Our graph notation is standard (for undefined terms, see West [34]). Let $L = \{l_1, \ldots, l_r\}$, where $l_1, \ldots, l_r$ are positive integers. An *L-bipartite covering* of a graph G is a collection of bicliques such that every edge of G is contained in exactly l_i bicliques for some $l_i \in L$. As each edge is a biclique, every graph has an L-biclique covering. We denote by $\text{bp}_L(G)$ the minimum number of bicliques in an L-bipartite covering of G. The parameter $\text{bp}_{\{1\}}(G)$ is *the biclique partition number* of G (see the survey [21]) and the Graham–Pollak Theorem states that $\text{bp}_{\{1\}}(K_n) = n - 1$. Also, $\text{bp}_{\{1,2,\ldots,t\}}(G)$, which will be denoted by $\text{bp}_t(G)$, is the *bipartite covering number* of order t or the *t-biclique covering number* of G (see Alon [2] and Huang–Sudakov [17]). For fixed $t \geq 2$, Alon [2] showed that $(1+o(1))(t!/2^t)^{1/t}n^{1/t} \leq \text{bp}_t(K_n) \leq (1+o(1))tn^{1/t}$. Huang and Sudakov [17] improved Alon's lower bound to $(1+o(1))(t!/2^{t-1})^{1/t}n^{1/t} \leq \text{bp}_t(K_n)$. The gap between the two bounds is still fairly large, even in the case $t = 2$. For $L = \{\lambda\}$, the parameter $\text{bp}_L(K_n)$ was studied by De Caen, Gregory, and Pritikin [13]. When L is the set of all odd numbers, the parameter $\text{bp}_L(K_n)$ was investigated by Radhakrishnan, Sen, and Vishwanathan [24].

2.1. *Lower bounds for* $\text{bp}_L(G)$ *and* $\text{bp}_L(K_n)$

In this subsection, we obtain two lower bounds for $\text{bp}_L(G)$ and $\text{bp}_L(K_n)$ using algebraic methods. Our first result, Theorem 2.1, gives a lower bound for $\text{bp}_L(G)$ for any list L and any graph G. Our proof is a modification of the proof of Huang and Sudakov [17] yielding a lower bound of $(1+o(1))(t!/2^{t-1})^{1/t}n^{1/t}$ for $\text{bp}_t(K_n)$. These authors follow a proof of the Graham–Pollak Theorem given by Peck [23], whereas we complete the argument in the spirit of a different proof of the Graham–Pollak Theorem, due to Witsenhausen (see, e.g. [8]). We believe that our proof can yield useful information about the case of equality. Our second result, Theorem 2.3, gives a lower bound for $\text{bp}_L(K_n)$ for any list L. Our proof is inspired by an argument of Alon [2]. We make a small improvement to Alon's proof (one that does not change the order of magnitude of the lower bound).

Theorem 2.1. *If there exists an L-bipartite covering of a graph, with d bicliques, then*

$$\max(n_+(A(G)), n_-(A(G))) \leq \sum_{l \in L} 2^{l-1}\binom{d}{l},$$

where $n_+(A(G))$ *and* $n_-(A(G))$ *are the numbers of positive and negative eigenvalues of* $A(G)$*, respectively.*

Proof. Suppose that the edges of G are covered by the bicliques $\{B(U_i, W_i)\}_{i=1}^{d}$ such that the number of times every edge in G is covered is in the list L. For every subset of indices $S \subset [d]$ with $|S| \in L$, let $H_S = \bigcap_{j\in S} B(U_j, W_j)$, and let A_S be its $n \times n$ adjacency matrix. If $|S| = s$, then for every $\{0, 1\}$ vector $z = (z_1, \ldots, z_{s-1})$ consider the complete bipartite graph with partite sets X_z and Y_z defined as follows:

$$X_z = \bigcap_{j:z_j=0} U_j \bigcap_{j:z_j=1} W_j \bigcap U_s \quad \text{and} \quad Y_z = \bigcap_{j:z_j=0} W_j \bigcap_{j:z_j=1} U_j \bigcap W_s.$$

As z ranges over all $\{0, 1\}$-vectors of length $s - 1$, these bicliques are edge-disjoint and their union is H_S. Thus, H_S is the disjoint union of at most 2^{s-1} bicliques, which means that A_S can be written as the sum of the adjacency matrices of at most 2^{s-1} bicliques.

Now we can write

$$A(G) = \sum_{k=1}^{r} c_k \sum_{S\subset[d],|S|=l_k} A_S, \tag{1}$$

where $c_i = 1 - \sum_{k=1}^{i-1} c_k \binom{l_i}{l_k}$. This implies that $A(G)$ is a linear combination of the $n \times n$ adjacency matrices of at most M bicliques, where $M = \sum_{l\in L} \binom{d}{l} 2^{l-1}$. Denote by u_i and v_i the characteristic vectors of the partite sets of the i-th biclique in this decomposition. If D_i is the $n \times n$ adjacency matrix of the i-th biclique, then $D_i = u_i v_i^T + v_i u_i^T$. Let

$$W = \text{Span}\{w \in \mathbb{R}^n : w^T u_i = 0, \text{ for } 1 \le i \le M\},$$
$$P = \text{Span}\{\text{Eigenvectors of the positive eigenvalues of } A\}.$$

Since W consists of n-dimensional vectors that are all orthogonal to M vectors, we have $\dim(W) \ge n - M$. Since $p^T Ap > 0$ for all nonzero $p \in P$, we have $W \cap P = \{0\}$. Therefore $\dim(W) \le n - \dim(P) = n - n_+(A(G))$. This implies $M = \sum_{l\in L} \binom{d}{l} 2^{l-1} \ge n_+(A(G))$. The inequality $M \ge n_-(A(G))$ can be proved similarly. □

As $n_+(A(K_n)) = 1$ and $n_-(A(K_n)) = n - 1$, the previous theorem implies the following result.

Corollary 2.2. *If K_n has an L-bipartite covering of size d, then* $n - 1 \le \sum_{l\in L} 2^{l-1} \binom{d}{l}$.

Theorem 2.3. *If K_n has an L-bipartite covering of size d, then* $n \le \sum_{i=0}^{|L|} 2^i \binom{d}{i} - \sum_{i=0}^{|L|-1} 2^i \binom{d-1}{i} = 2^{|L|} \binom{d}{|L|} + \sum_{i=1}^{|L|-1} 2^i \binom{d-1}{i-1}$.

Proof. Let $B_1, \ldots, B_d$ with partite sets $U_1, \ldots, U_d$ and $W_1, \ldots, W_d$ respectively, be complete bipartite graphs that form an L-bipartite covering of K_n with size d. Let $N = V(K_n) = \{1, \ldots, n\}$. For $i \in N$, define a polynomial $P_i(x_1, \ldots, x_d, y_1, \ldots, y_d)$ by

$$P_i(x_1, \ldots, x_d, y_1, \ldots, y_d) = \prod_{l\in L} \left(\sum_{p:i\in U_p} x_p + \sum_{q:i\in W_q} y_q - l \right).$$

For $i \in N$, we denote by e_i the vector $(w_{i1}, \ldots, w_{id}, u_{i1}, \ldots, u_{id})$, where $u_{ip} = 1$ if $i \in U_i$ and 0 otherwise, and $w_{iq} = 1$ if $i \in W_q$ and 0 otherwise. Now $\sum_{p:i\in U_p} x_p + \sum_{q:i\in W_q} y_q$ evaluated at e_j is precisely the number of times the edge between i and j appears in one of the bicliques $B_1, \ldots, B_d$. Thus

$$P_i(e_j) = 0 \quad \text{for all } 1 \le i \ne j \le n \quad \text{and} \quad P_i(e_i) = \prod_{l\in L}(-l) \ne 0.$$

Let $\overline{P}_i(x_1, \ldots, x_d, y_1, \ldots, y_d)$ be the multilinear polynomial obtained from P_i by replacing each monomial of the form $\prod_{s\in S} x_s^{\alpha_s} \prod_{t\in T} y_t^{\beta_t}$, where $\alpha_s \ge 1$ and $\beta_t \ge 1$ for each s and t, by the linear monomial $\prod_{s\in S} x_s \prod_{t\in T} y_t$. When $x_1, \ldots, x_d, y_1, \ldots, y_d \in \{0, 1\}$, we have $P_i(x_1, \ldots, x_d, y_1, \ldots, y_d) = \overline{P}_i(x_1, \ldots, x_d, y_1, \ldots, y_d)$ for $1 \le i \le n$. By the above equation, this implies $\overline{P}_i(e_j) = 0$ for $1 \le i \ne j \le n$ and $\overline{P}_i(e_i) \ne 0$ for $1 \le i \le n$. By the diagonal criterion (see [2,6] for more details), the multilinear polynomials $\overline{P_1}, \ldots, \overline{P_n}$ are linearly independent. Each $\overline{P_i}$ is a multilinear polynomial of degree at most $r = |L|$.

By definition, each $\overline{P}_i$ does not have any monomials that contain both x_j and y_j for the same j. Also, without loss of generality, we may assume that vertex 1 always goes in partite set U_i whenever vertex 1 appears in a biclique B_i. Thus, none of the $\overline{P_i}$'s contains y_1. The polynomials $\overline{P_i}$ are in the space generated by $\prod_{x\in S} x_s \prod_{t\in T} y_t$, where S and T range over all subsets of N satisfying $|S| + |T| \le r$ and $S \cap T = \emptyset$ and $1 \notin T$. Since there are $\sum_{i=0}^{r} 2^i \binom{d}{i} - \sum_{i=0}^{r-1} 2^i \binom{d-1}{i}$ pairs $\{S, T\}$, these facts imply $n \le \sum_{i=0}^{r} 2^i \binom{d}{i} - \sum_{i=0}^{r-1} 2^i \binom{d-1}{i}$. □

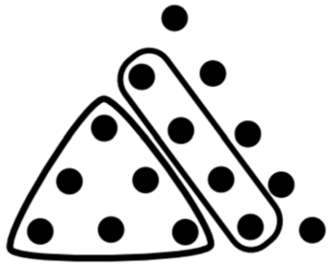

Fig. 1. A biclique in the triangular grid.

2.2. Constructive upper bounds for $\mathrm{bp}_L(K_n)$

In this subsection, we give examples of lists with $|L| > 1$ such that the previous lower bounds give the right order of magnitude for $\mathrm{bp}_L(K_n)$. Theorem 2.1 will yield better bounds for lists of the form $L = \{1, \ldots, t\}$, while Theorem 2.3 will be stronger in the other cases. We remark that there are examples of lists where neither of the results in the previous section produce the correct order of magnitude of $\mathrm{bp}_L(K_n)$. For example, if L contains only integers congruent to $a \pmod p$ for some prime number p and natural number a coprime with p, then one can show that $\mathrm{bp}_L(K_n) \geq \frac{n-1}{2}$ (see the remark below and also [24]), while Corollary 2.2 and Theorem 2.3 produce only sublinear lower bounds.

Remark 1. If K_n is covered by bicliques $B_1, \ldots, B_d$ so that each edge is covered a number of times congruent modulo p to a number not divisible by p, then

$$A(K_n) \equiv \sum_{i=1}^{d} a \cdot A(B_i) \pmod p.$$

This implies $n - 1 \leq \mathrm{rank}(A(K_n)) \leq \sum_{i=1}^{d} \mathrm{rank}(a \cdot A(B_i)) = 2d$, which implies $d \geq \frac{n-1}{2}$.

The upper bound in the first example is a construction by Alon [2], which we recall here for completeness. The upper bound in the second example is a modification of a different construction of Alon [2]. The third and fourth example seem to be new. The fifth example is due to Alon (private communication to the authors).

Example 1. $L = \{1, 2\}$.

Let $d = \mathrm{bp}_L(K_n)$. Corollary 2.2 gives $n \leq 1 + d^2$; that is, $d \geq \sqrt{n-1}$. For the upper bound, let $k = \lfloor \sqrt{n} \rfloor$. This means $k^2 \leq n < (k+1)^2$. If $n = k^2$, then we put the n vertices into a grid with k rows, $\{R_i\}_{i=1}^{k}$, and k columns, $\{C_i\}_{i=1}^{k}$. The $2k-2$ bicliques (described by their color classes) $\{(R_i, \cup_{j=i+1}^{k} R_j)\}_{i=1}^{k-1}$ and $\{(C_i, \cup_{j=i+1}^{k+1} C_j)\}_{i=1}^{k-1}$ form a $\{1, 2\}$-covering of K_n. Thus if $n = k^2$, then

$$\sqrt{n-1} \leq \mathrm{bp}_{\{1,2\}}(K_n) \leq 2\lfloor \sqrt{n} \rfloor - 2.$$

One can show by a similar argument that if $k^2 < n < k^2 + k$, then $\mathrm{bp}_{\{1,2\}}(K_n) \leq 2\lfloor \sqrt{n} \rfloor - 1$, and if $k^2 + k < n < (k+1)^2$, then $\mathrm{bp}_{\{1,2\}}(K_n) \leq 2\lfloor \sqrt{n} \rfloor$. By a case analysis, which we omit here, we have checked that $\mathrm{bp}_2(K_n)$ equals this upper bound for $3 \leq n \leq 9$ and $n = 12$.

Example 2. $L = \{2, 4, \ldots, 2k\}$.

Let $d = \mathrm{bp}_L(K_n)$. Theorem 2.3 gives

$$n \leq 2^k \binom{d}{k} + \sum_{i=1}^{k-1} 2^i \binom{d-1}{i-1}.$$

Thus d is bounded below by roughly $\frac{(k!n)^{1/k}}{2}$. To see the upper bound, consider the complete graph on n vertices, where $n = \binom{d}{k} + \binom{d}{k-2} + \cdots + \binom{d}{k-2\lfloor k/2 \rfloor}$. Index the vertices by the subsets of $[d]$ of sizes $k, k-2, k-4, \ldots, k-2\lfloor k/2 \rfloor$. Cover K_n with d bicliques, where in biclique i we put $A \in U_i$ if and only if $i \in A$ and $A \in W_i$ if and only if $i \notin A$. The number of edges joining A and B equals $|A \Delta B|$; hence the number of times each edge is covered is contained in $\{2, 4, \ldots, 2k\}$. Thus, d is bounded above by roughly $(k!n)^{1/k}$.

Example 3. $L = \{2, 3\}$.

Let $d = \mathrm{bp}_L(K_n)$. Theorem 2.3 gives $n \leq 2(d^2 - d + 1)$, which implies $d > \sqrt{\frac{n}{2}}$. For the upper bound, assume that $n = \binom{k}{2}$ and arrange the vertices of K_n in a triangular grid with k vertices on each side. For each direction parallel with one of the sides of the grid, we construct $k-1$ bicliques as follows. When the direction of the side is horizontal and $1 \leq i \leq k-1$, construct a biclique whose first partite set is the i-th row (from the bottom) of the triangular grid and second partite set is formed by the rows $i+1, \ldots, k$ (see Fig. 1). After a similar construction for the other two sides, we obtain $3(k-1)$ bicliques that cover the edges of K_n so that each edge is covered two or three times. If xy is an edge of K_n and the points corresponding to xy in the triangular grid create a line parallel to one of the sides of the triangular grid, then xy is covered exactly two times. Otherwise, xy is covered exactly three times.

Thus, we obtain $\frac{k}{2} < \mathrm{bp}_{\{2,3\}}(K_{\binom{k}{2}}) < 3k$.

Example 4. $L = \{1, 3\}$.

Remark 1 implies that $\mathrm{bp}_{1,3}(K_n) \geq \frac{n-1}{2}$ for every n (this lower bound will be $\frac{n}{2}$ when n is even). In [24], the authors show that actually $\mathrm{bp}_{\{1,3\}}(K_8) = 4$. Using this fact and a similar argument to one in [13] (which we will describe below), we can obtain the upper bound $\mathrm{bp}_{\{1,3\}}(K_n) \leq \frac{4n}{7} + 2$.

Proposition 2.4. *For any list L and natural numbers a, b.*

$$\mathrm{bp}_L(K_{a+b-1}) \leq \mathrm{bp}_L(K_a) + \mathrm{bp}_L(K_b).$$

Proof. Denote the vertex set of K_a by A and the vertex set of K_b by B. Identify one element of A with one element of B and call this element x. We will think of the vertex set of K_{a+b-1} as being $A \cup B$ with an overlap at x.

We will now construct an L-covering of K_{a+b-1} with $\mathrm{bp}_L(K_a)+\mathrm{bp}_L(K_b)$ bicliques. Leave the $\mathrm{bp}_L(K_a)$ bicliques in an optimal covering as they were. Change the $\mathrm{bp}_L(K_b)$ bicliques in an optimal covering of K_b into bicliques in K_{a+b-1} as follows:

1. if a biclique (U, V) does not contain x, then leave it unchanged.
2. if a biclique (U, V) contains x (say $x \in U$), then replace it by $(A \cup U, V)$.

The edges inside K_a are L-covered. The edges inside K_b are L-covered. Any edge pq with $p \in A \setminus \{x\}$ and $q \in B \setminus \{x\}$) is covered the same number of times as the edge xq in K_b. This finishes the proof. □

The previous result implies that $\mathrm{bp}_{\{1,3\}}(K_n) \leq \mathrm{bp}_{\{1,3\}}(K_{n-7}) + \mathrm{bp}_{\{1,3\}}(K_8) = \mathrm{bp}_{\{1,3\}}(K_{n-7}) + 4$, which gives $\mathrm{bp}_{\{1,3\}}(K_n) \leq \frac{4n}{7} + 2$. Thus, $\lfloor \frac{n}{2} \rfloor \leq \mathrm{bp}_{\{1,3\}}(K_n) \leq \frac{4n}{7} + 2$.

The following example was suggested to the authors by Noga Alon (private communication in October 2011).

Example 5. $L = \{1, 2, 4\}$.

Using the construction from Example 2 (for $k = 2$), we obtain $\mathrm{bp}_L(K_n) \leq (\sqrt{2}+o(1))n^{1/2}$. Let $t = \mathrm{bp}_L(K_n)$. If K_n is covered by bicliques $B_1, \ldots, B_t$ so that the number of times each edge is covered lies in L, then $A(B_1) + \cdots + A(B_t) \equiv A \mod 3$, where $A = (a_{i,j})_{1\leq i,j\leq n}$ is a symmetric $n \times n$ matrix having 0 diagonal and $a_{i,j} = \pm 1 \mod 3$ for $i \neq j$. The subadditivity of the rank implies $t \geq \mathrm{rank}(A)/2$. Since the matrix $(a_{i,j}^2)_{1\leq i,j\leq n}$ is the identity I_n, a result of Alon [3, Lemma 2.3] implies $\mathrm{rank}(A) \geq (\sqrt{2} - o(1))n^{1/2}$. Thus, $\mathrm{bp}_L(K_n) \geq \left(1/\sqrt{2} - o(1)\right) n^{1/2}$.

Example 6. $L = \{\lambda\}$.

When $L = \{\lambda\}$, De Caen, Gregory, and Pritikin conjectured [13] that $\mathrm{bp}_L(K_n) = n - 1$ for n large enough. This is known to be true for $\lambda \leq 18$ and is related to interesting problems in design theory and finite geometry (see [13] for more details).

To the knowledge of the authors, there are not many lists of constant size greater than 1 (not depending on n) for which the exact value of $\mathrm{bp}_L(K_n)$ is known. It is well known that if $L = \{1, \ldots, \lfloor \log_2 n \rfloor\}$, then $\mathrm{bp}_L(K_n) = \lceil \log_2 n \rceil$. In [24], the authors showed that $\mathrm{bp}_L(n) = \frac{n}{2}$ for infinitely many values of even n when L is the list of odd numbers less than n; these authors also proved similar results when L is the list of numbers congruent to 1 (mod p), p a fixed prime.

3. The Graham–Pollak theorem for hypergraphs

We first explain the hypergraph notation used in this section (see also [7]). Let $[n]$ denote the set $\{1, \ldots, n\}$ and $\binom{[n]}{r}$ denote the family of r-subsets of $[n]$. If $X_1, \ldots, X_r$ are disjoint subsets of $[n]$, then $\prod_{i=1}^{r} X_i$ denotes the family of r-subsets $\{x_1, \ldots, x_r\}$ where $x_i \in X_i$ for $1 \leq i \leq r$. The complete r-partite r-uniform hypergraph $[X_1, \ldots, X_r]$ with parts $X_1, \ldots, X_r$ is the r-uniform hypergraph with edge set $\prod_{i=1}^{r} X_i$.

Given an r-uniform hypergraph H, let $f_r(H)$ denote the minimum number of complete r-partite r-uniform hypergraphs needed to partition the edge set of H. In this section, we study $f_r(K_n^{(r)})$, writing it as $f_r(n)$. The Graham–Pollak Theorem states that $f_2(n) = \mathrm{bp}(K_n) = n - 1$.

Aharoni and Linial (cf. [1]) raised the natural problem of determining or approximating $f_r(n)$ for $r > 2$. In particular, they asked whether $f_r(n)$ is a nonlinear function of n for some fixed $r > 2$. Alon [1] answered this question and proved that $f_3(n) = n - 2$ and $f_r(n) = \Theta(n^{\lfloor r/2 \rfloor})$ for fixed $r \geq 4$ and $n \to \infty$. When $r \geq 4$, the exact value of $f_r(n)$ is known in very few cases.

The best known bounds for $f_r(n)$ were obtained by Cioabă, Küngden, and Verstraëte [9], who improved previous results of Alon [1] and showed that

$$\frac{2\binom{n-1}{k}}{\binom{2k}{k}} \le f_{2k}(n) \tag{2}$$

and

$$f_{2k}(n) \le f_{2k+1}(n+1) \le \binom{n-k}{k}. \tag{3}$$

In Section 3.1, we improve the above upper bound (3) for $f_{2k}(n)$ from [9] when $k \ge 3$. In Section 3.2, we determine the exact value of $f_r(r+2)$ whenever $r \ge 2$ and we asymptotically determine $f_r(r+t)$ for fixed t and $r \to \infty$. These results will show that neither of the bounds (2) and (3) gives the exact value of $f_r(n)$. This suggests that determining the exact value of $f_r(n)$ is a difficult problem.

3.1. Improved upper bound for $f_{2k}(n)$ when $k \ge 3$

In this subsection, we use a recursive construction to improve the upper bound $f_{2k}(n) \le \binom{n-k}{k}$ for $k \ge 3$. We obtain this improvement by considering the n vertices as the union of a set of j vertices and a set of $n-j$ vertices for some j. We can then use some upper bounds on the smaller sets from Section 3.2 (more precisely, the fact that $f_6(8) = 9$) to bound $f_{2k}(n)$. The following is the main result of this section.

Theorem 3.1. *For $n \ge 2k+2 \ge 8$,*

$$f_{2k}(n) \le \binom{n-k}{k} - 2\left\lfloor \frac{n}{16} \right\rfloor \binom{\left\lfloor \frac{n}{2} \right\rfloor - k + 3}{k-3}.$$

To prove this result, we will need Lemma 3.2 and Proposition 3.3, which we describe below.

Lemma 3.2. *Fix j, k. For $n \ge j$, we have*

$$\binom{n-k}{k} = \sum_{i=0}^{2k} \binom{j - \left\lceil \frac{i}{2} \right\rceil}{\left\lfloor \frac{i}{2} \right\rfloor} \binom{n-j-\left\lceil \frac{2k-i}{2} \right\rceil}{\left\lfloor \frac{2k-i}{2} \right\rfloor}.$$

Proof. We use induction on n. Since $\binom{m}{t} = 0$ for $m < 0$ and $\binom{0}{0} = 1$, both sides equal $\binom{j-k}{k}$ when $n = j$. For larger n, applying Pascal's Identity to the second factor in the sum breaks the sum into two sums that, by the induction hypothesis, equal $\binom{(n-1)-k}{k}$ and $\binom{(n-2)-(k-1)}{k}$, which sum to $\binom{n-k}{k}$. □

Proposition 3.3. *For $n \ge 8$,*

$$f_6(n) \le \binom{n-3}{3} - \left\lfloor \frac{n}{8} \right\rfloor.$$

Proof. Let $n = 8j + r$ with $0 \le r < 8$. We use induction on j. For the base case, we have $n = 8 + r$, where $0 \le r < 8$. We break the n vertices into one part of size r and one part of size 8. A decomposition of $K_n^{(6)}$ into complete 6-partite 6-uniform hypergraphs can be obtained from a partition of $K_8^{(i)}$ into $f_i(8)$ complete i-partite i-uniform hypergraphs and a partition of $K_r^{(6-i)}$ into $f_{6-i}(r)$ complete $(6-i)$-partite $(6-i)$-uniform hypergraphs, when i takes values between 0 and 6. Thus

$$f_6(n) \le \sum_{i=0}^{6} f_i(8) f_{6-i}(r).$$

We bound everything above using (3) except for the term $f_6(8)$, where we use Theorem 3.7. Using Lemma 3.2, this proves the base case.

For $j > 1$, we provide a recursive construction by breaking the n vertices into a set of size 8 and a set of size $n-8$. We find

$$f_6(n) \le \sum_{i=0}^{6} f_i(8) f_{6-i}(n-8) = f_6(8) + f_6(n-8) + \sum_{i=1}^{5} f_i(8) f_{6-i}(n-8).$$

For i from 1 to 5 we bound above using (3), and for $i = 0$ and $i = 6$ we use the inductive hypothesis and base case. Applying Lemma 3.2 gives

$$f_6(n) \le \binom{n-3}{3} - 1 - \left\lfloor \frac{n-8}{8} \right\rfloor,$$

which proves the lemma. □

We are now ready to describe the proof of Theorem 3.1.

Proof. It suffices to prove the claim for even n, so we assume this for convenience. Again, a decomposition of $K_n^{(2k)}$ into complete $2k$-partite $2k$-uniform hypergraphs can be obtained from a decomposition of $K_{\frac{n}{2}}^{(i)}$ into $f_i(\frac{n}{2})$ complete i-partite i-uniform subgraphs and a partition of $K_{\frac{n}{2}}^{(2k-i)}$ into $f_{2k-i}(\frac{n}{2})$ complete $(2k-i)$-partite $(2k-i)$-uniform hypergraphs, when i takes values from 0 to $2k$. Thus

$$f_{2k}(n) \le 2f_6\left(\frac{n}{2}\right) f_{2k-6}\left(\frac{n}{2}\right) + \sum_{i \ne 6, 2k-6} f_i\left(\frac{n}{2}\right) f_{2k-i}\left(\frac{n}{2}\right).$$

We bound each term of the previous sum from above by (3) for all $f_j\left(\frac{n}{2}\right)$ except for $i = 6$, where we use the bound given by Proposition 3.3. Using Lemma 3.2, we obtain the desired result:

$$f_{2k}(n) \le \binom{n-k}{k} - 2\left\lfloor \frac{n}{16} \right\rfloor \binom{\left\lfloor \frac{n}{2} \right\rfloor - k + 3}{k-3}. \quad \square$$

3.2. *Determining $f_r(n)$ when $n - r$ is a constant*

In this section, we determine the exact value of $f_r(r+2)$. Exact values for $f_r(n)$ are only known for $r \in \{2, 3\}$. We also determine the asymptotics of $f_r(r+t)$ when t is fixed and $r \to \infty$.

When $n = r+2$, the complement of each hyperedge can be seen as an edge of K_{r+2}, and we refer to K_{r+2} as the *complement* of $K_{r+2}^{(r)}$. If we decompose $K_{r+2}^{(r)}$ into complete r-partite r-uniform subgraphs, then the complements of the hyperedges will decompose the graph K_{r+2}. A complete r-partite r-uniform subhypergraph of $K_{r+2}^{(r)}$ has one of the following forms:

1. It has r partite sets of size 1, and it produces one edge in the complement.
2. It has $r-1$ partite sets of size 1 and one partite set of size 2, and it produces a path of two edges in the complement.
3. It has $r-1$ partite sets of size 1 and one set of size 3, and it produces a triangle in the complement.
4. It has $r-2$ partite sets of size 1 and two partite sets of size 2, and it produces a 4-cycle in the complement.

Thus, partitioning $K_{r+2}^{(r)}$ into complete r-partite r-uniform hypergraphs is equivalent to decomposing K_{r+2} into copies of $K_2, K_{1,2}, K_3$, and C_4.

Proposition 3.4. *For any natural number r, $f_{8r-1}(8r+1) = r(8r+1)$.*

Proof. Because $4 | 4r(8r+1) = \binom{8r+1}{2}$, one can partition K_{8r+1} into $r(8r+1)$ copies of C_4 (see Sajna [26]). The lower bound $f_{8r-1}(8r+1) \ge \frac{\binom{8r+1}{2}}{4} = r(8r+1)$ holds since each of K_2, K_3, C_4, or $K_{1,2}$ contains at most four edges. □

Proposition 3.5. *For $k \ge 2$, $f_{2k}(2k+2) \ge \lceil \frac{(k+1)(2k+3)}{4} \rceil$.*

Proof. Consider an optimal decomposition of K_{2k+2} using $\{K_2, K_{1,2}, K_3, C_4\}$. Since each vertex of K_{2k+2} has odd degree, each vertex must be incident with at least one copy of K_2 or $K_{1,2}$. Suppose that t graphs with at most two edges have been used, where $t \ge k+1$. At least $\binom{2k+2}{2} - 2t$ edges remain uncovered, and at least $\frac{\binom{2k+2}{2}-2t}{4}$ graphs from K_3 or C_4 must be used to cover these edges. Thus, $f_{2k}(2k+2) \ge t + \frac{\binom{2k+2}{2}-2t}{4} \ge \frac{\binom{2k+2}{2}+2(k+1)}{4} = \frac{(k+1)(2k+3)}{4}$. □

Proposition 3.6. *For any natural number r, $f_{8r}(8r+2) = 8r^2 + 5r + 1$.*

Proof. Let v be a vertex of $K_{8r+2}^{(8r)}$. The hyperedges containing v can be partitioned into $f_{8r-1}(8r+1)$ hypergraphs. The hyperedges not containing v can be partitioned into $f_{8r}(8r+1)$ hypergraphs. This gives $f_{8r}(8r+2) \le f_{8r-1}(8r+1) + f_{8r}(8r+1)$. The hypergraph $K_{8r+1}^{(8r)}$ can be partitioned into the following $4r+1$ complete $8r$-partite $8r$-uniform subgraphs: $G_1 = \{\{1,2\},\{3\},\{4\},\dots,\{8r+1\}\}$, $G_2 = \{\{1\},\{2\},\{3,4\},\dots,\{8r+1\}\}, \dots, G_{4r} = \{\{1\},\{2\},\dots\{8r-1,8r\},\{8r+1\}\}$, $G_{4r+1} = \{\{1\},\{2\},\dots,\{8r\}\}$. Thus $f_{8r}(8r+1) \le 4r+1$. Proposition 3.4 gives $f_{8r-1}(8r+1) = 8r^2 + r$. Combining all these facts, we get $f_{8r}(8r+2) \le 8r^2 + 5r + 1$. The lower bound is given by Proposition 3.5. □

Theorem 3.7. *For $k \geq 2$, $f_{2k}(2k+2) = f_{2k+1}(2k+3) = \lceil \frac{2k^2+5k+3}{4} \rceil$.*

Proof. We prove that $f_{2k}(2k+2) = \lceil \frac{2k^2+5k+3}{4} \rceil$ by induction on k. For the base cases, it is known that $f_2(4) = 3$. Eq. (3) gives $f_4(6) \leq 6$, which is the lower bound given by Proposition 3.5. We also have $f_6(8) \leq f_7(9) = 9$ by Proposition 3.4 and $f_6(8) \geq 9$ from Proposition 3.5. Assume that $f_{2k}(2k+2) = \lceil \frac{2k^2+5k+3}{4} \rceil$. Proposition 3.5 implies $f_{2k+2}(2k+4) \geq \lceil \frac{2k^2+9k+10}{4} \rceil$. Consider the following decomposition of K_{2k+4} using $\{K_2, K_{1,2}, K_3, C_4\}$. Pick any two vertices of K_{2k+4}. Take an optimal decomposition into $\left\lceil \frac{2k^2+5k+3}{4} \right\rceil$ copies of $K_2, K_{1,2}, K_3$, or C_4 of the complete graph induced by the other $2k+2$ vertices. To complete the decomposition, we need to cover all edges joining the $2k+2$ original vertices to the two remaining vertices, plus the edge joining the two remaining vertices. We use a copy of K_2 to cover the edge joining the two vertices, and $k+1$ copies of C_4, each using these two vertices and two of the other vertices. Thus, we can decompose K_{2k+4} into copies of $K_2, K_{1,2}, K_3$, and C_4 using $\lceil \frac{2k^2+5k+3}{4} \rceil + 1 + (k+1) = \lceil \frac{2k^2+9k+11}{4} \rceil$ subgraphs. Thus, $\lceil \frac{2k^2+9k+10}{4} \rceil \leq f_{2k+2}(2k+4) \leq \lceil \frac{2k^2+9k+11}{4} \rceil$.

These bounds are the same whenever $k \equiv 0, 1 \pmod 4$. If $k \equiv 3 \pmod 4$, then the result follows from Proposition 3.6. If $k \equiv 2 \pmod 4$, then we want to find $f_{8r+6}(8r+8)$ for some r. However, $f_{8r+6}(8r+8) \leq f_{8r+7}(8r+9) = (r+1)(8r+9)$ by Proposition 3.4. This completes our proof that $f_{2k}(2k+2) = \lceil \frac{2k^2+5k+3}{4} \rceil$. We know that $f_{2k+1}(2k+3) \geq f_{2k}(2k+2) = \lceil \frac{2k^2+5k+3}{4} \rceil$ by Eq. (3). To prove equality, one can use induction on k and a similar construction to the one above. We omit the details. □

The following result determines $f_r(r+t)$ asymptotically for fixed t and large r.

Proposition 3.8. *If $t \geq 3$, then $f_{2k}(2k+t) \sim \frac{k^t}{t!}$ and $f_{2k+1}(2k+t) \sim \frac{k^{t-1}}{(t-1)!}$ as $k \to \infty$.*

Proof. Let k and t be integers with $k \geq t \geq 3$. Since a complete $2k$-partite $2k$-uniform hypergraph on $2k+t$ vertices has at most 2^t hyperedges, we obtain $f_{2k}(n) \geq \frac{\binom{2k+t}{2k}}{2^t} = \frac{\binom{2k+t}{t}}{2^t}$. The upper bound (3) gives $f_{2k}(2k+t) \leq \binom{k+t}{k}$, and this yields $f_{2k}(2k+t) \sim \frac{k^t}{t!}$ as $k \to \infty$. The proof of $f_{2k+1}(2k+t) \sim \frac{k^{t-1}}{(t-1)!}$ is similar and is omitted. □

4. The Alon–Saks–Seymour conjecture

Alon, Saks, and Seymour (cf. Kahn [18]) conjectured that $\chi(G) \leq \text{bp}(G) + 1$ for every graph G, where $\chi(G)$ is the chromatic number of G. The Alon–Saks–Seymour Conjecture can be seen as a more general statement of the Graham–Pollak Theorem, since $\chi(K_n) = n$ and $\text{bp}(K_n) = n - 1$. Recently, Huang and Sudakov [17] disproved the Alon–Saks–Seymour Conjecture by constructing graphs G with arbitrarily large biclique partition number such that $\chi(G) > c(\text{bp}(G))^{6/5}$. Huang and Sudakov [17] conjectured that there exists a graph G with biclique partition number k and chromatic number at least $2^{c \log^2 k}$, for some constant positive constant c, and this problem remains open.

The Huang–Sudakov construction was generalized in [10], where the authors constructed graphs $G(n, k, r)$ on $n^{2k+2r+1}$ vertices with chromatic number greater than or equal to $\Omega(n^{2k+2r})$ and biclique partition number at most $O(n^{2k+2r-1})$, for n, k, r with $n \geq 2$, $k \geq 2$, and $r \geq 1$. The graphs $G(n, k, r)$ are also counterexamples to the previously-disproved Rank-Coloring Conjecture (see [10,17,31]) which stated that $\chi(G) \leq \text{rank}(A(G))$, where $A(G)$ is the adjacency matrix of G. The authors obtained asymptotically tight bounds for the ranks and the biclique partition numbers of these graphs in [10] (see also [27]). In this section, we extend these results, and we determine the exact value of the rank of the adjacency matrix of $G(n, k, r)$. We also compute the eigenvalues (and their multiplicities) of $G(2, k, r)$. These results imply that the order of magnitude of the biclique partition number of $G(n, k, r)$ is $\Theta(n^{2k+2r-1})$ for fixed k and r with $k \geq 2$ and $r \geq 1$.

Let Q_n be the n-dimensional cube with vertex set $\{0, 1\}^n$, where two vertices are adjacent if and only if they differ in exactly one coordinate. A *k-dimensional subcube* of Q_n is a subgraph of Q_n induced by a vertex subset of the following form

$$\{x \in Q_n : x_i = b_i, \text{ for } i \in B\},$$

where B is a set of $n - k$ fixed coordinates and each $b_i \in \{0, 1\}$. We represent the all ones and all zeros vectors as 1^n and 0^n respectively, and we define $Q_n^- = Q_n \setminus \{1^n, 0^n\}$.

For n, k, r with $n \geq 2$, $k \geq 1$, and $r \geq 1$, we define the graph $G(n, k, r)$ as follows. Its vertex set is

$$V(G(n, k, r)) = [n]^{2k+2r+1} = \{(x_1, \ldots, x_{2k+2r+1}) : x_i \in [n], \text{ for } 1 \leq i \leq 2k+2r+1\}.$$

For any two vertices x and y, let

$$\rho(x, y) = (\rho_1(x, y), \ldots, \rho_{2k+2r+1}(x, y)) \in \{0, 1\}^{2k+2r+1},$$

where $\rho_i(x, y) = 1$ if $x_i \neq y_i$ and $\rho_i(x, y) = 0$ if $x_i = y_i$.

We define adjacency in $G(n, k, r)$ as follows: the vertices x and y are adjacent in $G(n, k, r)$ if and only if $\rho(x, y) \in S$ where

$$S = Q_{2k+2r+1} \setminus [(1^{2k} \times Q_{2r+1}^-) \cup \{0^{2k} \times 0^{2r+1}\} \cup \{0^{2k} \times 1^{2r+1}\}].$$

The main results of this section are:

Theorem 4.1. *For n, k, r with $n > 2, k \geq 2, r \geq 1$, the rank of the adjacency matrix of $G(n, k, r)$ is $n^{2k+2r+1} - (n-1)^{2k}n^{2r+1} - n^{2k}(n-1)^{2r+1} + (n-1)^{2k+2r+1} + (n-1)^{2k}$.*

and

Proposition 4.2. *The spectrum of $G(2, k, r)$ is*

$$\begin{pmatrix} 2^{2k+2r+1} - 2^{2r+1} & 2^{2r+1} - 4 & 0 & -4 & -2^{2r+1} \\ 1 & 2^{2k-1} & 2^{2k+2r} + 2^{2k+2r-1} - 2^{2k-1} & 2^{2k+2r-1} - 2^{2k-1} & 2^{2k-1} - 1 \end{pmatrix}.$$

Next, we determine the rank of the adjacency matrix of $G(n, k, r)$. We will use the following graph operation called NEPS (Non-complete Extended P-Sum) introduced by Cvetković in his thesis [11] (see [12, p. 66] for more details including an explanation of the notation NEPS). The NEPS operation is a generalization of various other graph products including the Cartesian or Kronecker product of graphs.

Definition 1. Given $\mathcal{B} \subset \{0, 1\}^t \setminus \{0^t\}$ and graphs $G_1, \ldots, G_t$, the NEPS with basis $\mathcal{B}$ of the graphs $G_1, \ldots, G_t$ is the graph with vertex set $V(G_1) \times \cdots \times V(G_t)$ in which two vertices $(x_1, \ldots, x_t)$ and $(y_1, \ldots, y_t)$ are adjacent if and only if there is a t-tuple $(b_1, \ldots, b_t)$ in $\mathcal{B}$ such that $x_i = y_i$ exactly when $b_i = 0$ and $x_iy_i \in E(G_i)$ exactly when $b_i = 1$.

When $G_1, \ldots, G_t$ all are isomorphic to K_n, the NEPS with basis $\mathcal{B}$ of $G_1, \ldots, G_t$ is the graph whose vertex set is $[n]^t$ with $x \sim y$ if and only if $\rho(x, y) = (b_1, \ldots, b_t)$ for some $(b_1, \ldots, b_t) \in \mathcal{B}$. Hence, the graph $G(n, k, r)$ is the NEPS of $2k + 2r + 1$ copies of K_n with basis

$$S = Q_{2k+2r+1} \setminus [(1^{2k} \times Q^-_{2r+1}) \cup \{0^{2k} \times 0^{2r+1}\} \cup \{0^{2k} \times 1^{2r+1}\}].$$

Another important observation (see [12, Theorem 2.21, p. 68]) is that the adjacency matrix of the NEPS with basis $\mathcal{B}$ of $G_1, \ldots, G_t$ equals

$$\sum_{(b_1,\ldots,b_t)\in\mathcal{B}} A(G_1)^{b_1} \otimes \cdots \otimes A(G_t)^{b_t},$$

where $X \otimes Y$ denotes the Kronecker product of two matrices X and Y.

We are ready to complete the proof of Theorem 4.1.

Proof. The spectrum of a graph X will be denoted by Spec(X). Since $G = G(n, k, r)$ can be written as a NEPS of $2k + 2r + 1$ copies of the complete graph with set S, the spectrum of G (see [12, Theorem 2.21]) is given by

$$\text{Spec}(G) = \{f(\lambda_1, \ldots, \lambda_{2k+2r+1}) : \lambda_i \in \text{Spec}(K_n)\},$$

where

$$f(x_1, \ldots, x_{2k+2r+1}) = \sum_{(s_1,\ldots,s_{2k+2r+1})\in S} \prod_{i=1}^{2k+2r+1} x_i^{s_i}. \tag{4}$$

This can be written as

$$f(x_1, \ldots, x_{2k+2r+1}) = \prod_{i=1}^{2k+2r+1}(1 + x_i) - 1 - \prod_{i=1}^{2k} x_i \left(\prod_{i=2k+1}^{2k+2r+1}(1 + x_i) - 1 - \prod_{i=2k+1}^{2k+2r+1} x_i\right) - \prod_{i=2k+1}^{2k+2r+1} x_i \tag{5}$$

where each x_i runs through the spectrum of K_n. The eigenvalues of K_n are -1 with multiplicity $n-1$ and $n-1$ with multiplicity 1. Let a be the number of copies of -1 in the first $2k$ entries of $(x_1, \ldots, x_{2k+2r+1})$, and b the number of copies of -1 in the last $2r + 1$ entries of $(x_1, \ldots, x_{2k+2r+1})$. Let

$$A = \prod_{i=1}^{2k} x_i = (-1)^a(n-1)^{2k-a}$$

and

$$B = \prod_{2k+1}^{2k+2r+1} x_i = (-1)^b(n-1)^{2r+1-b}.$$

If $a = 0$ and $b = 0$, then we obtain the degree of regularity $n^{2k+2r+1} - 1 - (n-1)^{2k}[n^{2r+1} - 1 - (n-1)^{2r+1}] - (n-1)^{2r+1}$ as an eigenvalue with multiplicity 1.

If $a = 0$ and $b \neq 0$, then we obtain

$$-1 + (n-1)^{2k}(1+B) - B \quad (6)$$

as an eigenvalue with multiplicity $\binom{2r+1}{b}(n-1)^b$.

If $a \neq 0$ and $b = 0$, then we obtain

$$-1 - A(n^{2r+1} - 1 - (n-1)^{2r+1}) - (n-1)^{2r+1} \quad (7)$$

as an eigenvalue with multiplicity $\binom{2k}{a}(n-1)^a$.

Finally, if $a, b \neq 0$, then we obtain

$$A + AB - B - 1 \quad (8)$$

as an eigenvalue with multiplicity $\binom{2k}{a}\binom{2r+1}{b}(n-1)^{a+b}$.

We remark that A and B are powers of $n-1$ multiplied by positive or negative 1, so that in (6)–(8), the eigenvalues are -1 plus or minus some powers of $(n-1)$. Since $n > 2$, this means that we can only have 0 as an eigenvalue if these powers of $n-1$ add to 1. This puts many restrictions on A and B and thus on a and b described above.

In (6), the only way we can obtain 0 as an eigenvalue is if $B = -1$. This means that $b = 2r+1$, which corresponds to the situation in (5) where the first $2k$ positions contain all $n-1$, and all of the last $2r+1$ positions are -1. This gives eigenvalue 0 with multiplicity $(n-1)^{2r+1}$.

In (7) we cannot obtain 0 as an eigenvalue.

In (8) we obtain 0 as an eigenvalue if $A = 1$ or if $B = -1$. This happens when in (5), either the first $2k$ positions are -1 or the last $2r+1$ positions are -1. In this case, we obtain 0 as an eigenvalue with multiplicity $(n-1)^{2k}n^{2r+1} + n^{2k}(n-1)^{2r+1} - (n-1)^{2k+2r+1} - (n-1)^{2k} - (n-1)^{2r+1}$.

Thus, the multiplicity of eigenvalue 0 is

$$(n-1)^{2k}n^{2r+1} + n^{2k}(n-1)^{2r+1} - (n-1)^{2k+2r+1} - (n-1)^{2k},$$

which means the rank of $G(n, k, r)$ equals

$$n^{2k+2r+1} - (n-1)^{2k}n^{2r+1} - n^{2k}(n-1)^{2r+1} + (n-1)^{2k+2r+1} + (n-1)^{2k}. \quad \square$$

We remark here that in [10, p. 7] before Eq. (15), one should add *and when not all the last* $2r+1$ *positions are* $n-1$ in order for the Eq. (15) to hold. We thank Robert Coulter for observing this error, which does not affect the main result of [10]. We give below a short proof of Proposition 4.2.

Proof. The spectrum of $G(2, k, r)$ can be obtained by plugging the eigenvalues of K_n into the formula (5). Another way to compute the eigenvalues of $G(2, k, r)$ is by computing the eigenvalues of the complement G^c of $G(2, k, r)$. The complement of $G(n, k, r)$ is the Cayley graph of the additive group $\mathbb{F}_2^{2k+2r+1}$ with generating set $(1^{2k} \times Q_{2r+1}^-) \cup \{0^{2k} \times 1^{2r+1}\}$ (see [5,20] or [8] for details on calculating eigenvalues of Cayley graphs). We can express the $2^{2k+2r+1}$ eigenvalues of G^c as

$$\lambda_w = \sum_{t \in T} (-1)^{t \cdot w}$$

for $w \in \mathbb{F}_2^{2k+2r+1}$ and $T = \{(1^{2k} \times Q_{2r+1}^-) \cup \{0^{2k} \times 1^{2r+1}\}\}$.

Suppose that w has a1's in the first $2k$ positions and b1's in the last $2r+1$ positions, and denote by $w' = (w_{2k+1}, \ldots, w_{2k+2r+1})$ the projection of w onto the last $2r+1$ components.

If $b = 0$, then we obtain $\lambda_w = 1 + (-1)^a|Q_{2r+1}^-| = 1 + (-1)^a(2^{2r+1} - 2)$. Thus, we get eigenvalue $2^{2r+1} - 1$ with multiplicity 2^{2k-1} (corresponding to a even in the previous expression) and eigenvalue $-2^{2r+1} + 3$ with multiplicity 2^{2k-1} (corresponding to a odd).

If $b \neq 0$, then we have

$$\begin{aligned}
\lambda_w &= (-1)^b + (-1)^a \sum_{t \in Q_{2r+1}^-} (-1)^{t \cdot w'} \\
&= (-1)^b + (-1)^a \left(\sum_{t \in Q_{2r+1}} (-1)^{t \cdot w'} - (-1)^0 - (-1)^b \right) \\
&= (-1)^b + (-1)^a \left(0 - 1 - (-1)^b\right) \\
&= (-1)^b + (-1)^{a+1} + (-1)^{a+b+1}.
\end{aligned}$$

In this case, we obtain eigenvalue 3 with multiplicity $2^{2k+2r-1} - 2^{2k-1}$ (corresponding to a odd and b even) and eigenvalue -1 with multiplicity $2^{2k+2r} + 2^{2k+2r-1} - 2^{2k-1}$ (corresponding to the remaining cases). Thus, the spectrum of G^c is given by

$2^{2r+1}-1, 3, -1, 3-2^{2r+1}$, with multiplicities $2^{2k-1}, 2^{2k+2r-1}-2^{2k-1}, 2^{2r+2k}+2^{2r+2k-1}-2^{2k-1}, 2^{2k-1}$, respectively. By standard results in graph spectra (see [8,14]), we obtain the spectrum of $G(2, k, r)$ below (the first row denotes the distinct eigenvalues of $G(2, k, r)$, and the second row denotes their multiplicities):

$$\begin{pmatrix} 2^{2k+2r+1}-2^{2r+1} & 2^{2r+1}-4 & 0 & -4 & -2^{2r+1} \\ 1 & 2^{2k-1} & 2^{2k+2r}+2^{2k+2r-1}-2^{2k-1} & 2^{2k+2r-1}-2^{2k-1} & 2^{2k-1}-1 \end{pmatrix}. \quad \square$$

5. Conclusions

In this paper, we studied several variations of the Graham–Pollak Theorem.

We discussed a generalization of the Graham–Pollak Theorem to L-coverings of the complete graph. We obtained some lower bounds for $\mathrm{bp}_L(K_n)$ and $\mathrm{bp}_L(G)$ and presented several constructions determining the correct order of magnitude for $\mathrm{bp}_L(K_n)$ for several lists L. Our results motivate the following natural questions.

Open Problem 1. *For what lists L of constant size greater than one can the exact value for $\mathrm{bp}_L(K_n)$ be found for n large? Is it true that for any fixed list L, there exist constants c_L and q_L such that $\lim \frac{\mathrm{bp}_L(K_n)}{n^{1/q_L}} \to c_L$ as $n \to \infty$?*

Based on our results for small values of n, we pose the following questions:

Open Problem 2. *Is the upper bound given by the construction of Alon [2] and described in Example 1 the true value of $\mathrm{bp}_2(K_n)$? Does the limit $\lim_{n\to\infty} \frac{\mathrm{bp}_2(K_n)}{n^{1/2}}$ exist?*

We also discussed a generalization of the Graham–Pollak Theorem for uniform hypergraphs. With $f_r(n)$ denoting the minimum number of complete r-partite r-uniform hypergraphs necessary to partition the edges of the complete r-graph, we showed that

$$f_{2k}(n) \le \binom{n-k}{k} - 2\left\lfloor \frac{n}{16} \right\rfloor \binom{\left\lfloor \frac{n}{2} \right\rfloor - k + 3}{k-3}.$$

We also determined the exact value of $f_r(r+2)$ and the asymptotic value of $f_r(r+t)$ for t fixed and r large. We note here that we could not improve the upper bound on $f_{2k+1}(n)$ from [9] by a similar method, because Lemma 3.2 breaks down in this case. Thus the best upper bound known, $f_{2k+1}(n) \le \binom{n-k-1}{k}$, is still given by (3). A better upper bound for $f_{2k+1}(n)$ may still be found, since $f_{2k+1}(2k+3) = \lceil \frac{2k^2+5k+3}{4} \rceil$.

For fixed k and large n, the best known bounds for $f_{2k}(n)$ and $f_{2k+1}(n)$ are still far apart. We raise the following natural questions.

Open Problem 3. *For $r \ge 4$, is there a constant c_r such that $\frac{f_r(n)}{n^{\lfloor r/2 \rfloor}} \to c_r$ as $n \to \infty$? Is it true that $f_{2k}(n) \sim f_{2k+1}(n+1)$ for k fixed and $n \to +\infty$?*

We found the exact rank of the counterexamples to the Alon–Saks–Seymour Conjecture described in [10,17]. This extends the work from [10] and gives the exact order of magnitude for the biclique partition of these counterexamples. At this time, we do not know the exact order of magnitude of the chromatic number of these graphs, and this is a problem that might be worth studying. Another interesting open problem is constructing other graphs G with a large gap between $\chi(G)$ and $\mathrm{bp}(G)$.

Acknowledgments

This work was partially supported by a grant from the Simons Foundation (#209309 to Sebastian M. Cioabă) and by an NSA Young Investigator Grant. Part of this research was done while Michael Tait was a M.Sc. Student at the University of Delaware and was supported by a Unidel Fellowship. The authors thank Noga Alon, Andries Brouwer, Fan Chung, Ron Graham, André Kündgen and Jacques Verstraëte for interesting discussions. The authors are grateful to the two anonymous referees and Douglas West for their excellent comments and suggestions that have greatly improved the first version of this paper.

References

[1] N. Alon, Decompositions of the complete r-graph into complete r-partite r-graphs, Graphs Combin. 2 (1986) 95–100.
[2] N. Alon, Neighborly families of boxes and bipartite coverings, Algorithms Combin. 14 (1997) 27–31.
[3] N. Alon, Perturbed identity matrices have high rank: proof and applications, Combin. Probab. Comput. 18 (1–2) (2009) 3–15.
[4] N. Alon, P. Seymour, A counterexample to the rank-coloring conjecture, J. Graph Theory 13 (4) (1989) 523–525.
[5] L. Babai, Spectra of Cayley graphs, J. Combin. Theory Ser. B 27 (2) (1979) 180–189.

[6] L. Babai, P. Frankl, Linear Algebraic Methods in Combinatorics, in: Lecture Notes, University of Chicago.
[7] B. Bollobás, Combinatorics, Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability, Cambridge University Press, 1986.
[8] A.E. Brouwer, W.H. Haemers, Spectra of Graphs, in: Universitext, Springer, 2012, p. 250.
[9] S.M. Cioabă, A. Küngden, J. Verstraëte, On decompositions of complete hypergraphs, J. Combin. Theory Ser. A 116 (2009) 1232–1234.
[10] S.M. Cioabă, M. Tait, More counterexamples to the Alon–Saks–Seymour and rank-coloring conjectures, Electron. J. Combin. 18 (1–9) (2011) 26.
[11] D. Cvetković, Graphs and their spectra (Grafovi i njihovi spektri), Thesis, Univ. Beograd Publ. Elektrotehn. Fak. Ser. Math. Fiz. No. 354–No.356, 1971, pp. 1–50.
[12] D.M. Cvetković, M. Doob, H. Sachs, Spectra of Graphs: Theory and Application, in: Pure and Applied Mathematics, vol. 87, Academic Press, New York, London, 1980, p. 368.
[13] D. de Caen, D.A. Gregory, D. Pritikin, Minimum biclique partitions of the complete multigraph and related designs, in: Graphs, Matrices, and Designs, in: Lecture Notes in Pure and Appl. Math., vol. 139, Dekker, New York, 1993, pp. 93–119.
[14] C. Godsil, G. Royle, Algebraic Graph Theory, in: Springer Graduate Texts in Mathematics, vol. 207, 2001.
[15] R.L. Graham, H.O. Pollak, On the addressing problem for loop switching, Bell Syst. Tech. J. 50 (8) (1971) 2495–2519.
[16] R.L. Graham, H.O. Pollak, On embedding graphs in squashed cubes, in: Graph Theory and Applications, in: Lecture Notes in Math., vol. 303, Springer, Berlin, 1972, pp. 99–110.
[17] H. Huang, B. Sudakov, A counterexample to the Alon–Saks–Seymour conjecture and related problems, Combinatorica 32 (2012) 205–219.
[18] J. Kahn, Recent results on some not-so-recent hypergraph matching and covering problems, in: Extremal Problems for Finite Sets (1991), in: Bolyai Soc. Math. Stud., vol. 3, János Bolyai Math. Soc., Budapest, 1994, pp. 305–353.
[19] T. Kratzke, B. Reznick, D. West, Eigensharp graphs: decomposition into complete bipartite subgraphs, Trans. Amer. Math. Soc. 308 (1988) 637–653.
[20] L. Lovász, Spectra of graphs with transitive groups, Period. Math. Hungar. 6 (2) (1975) 191–195.
[21] S. Monson, N.J. Pullman, R. Rees, A survey of clique and biclique coverings and factorizations of (0,1)-matrices, Bull. Inst. Combin. Appl. 14 (1995) 17–86.
[22] N. Nisan, A. Wigderson, On rank vs. communication complexity, Combinatorica 15 (1995) 557–565.
[23] G.W. Peck, A new proof of a theorem of Graham and Pollak, Discrete Math. 49 (1984) 327–328.
[24] J. Radhakrishnan, P. Sen, S. Vishwanathan, Depth-3 arithmetic for $S_n^2(X)$ and extensions of the Graham–Pollack theorem, in: FST TCS 2000: Foundations of Software Technology and Theoretical Computer Science, New Delhi, in: Lecture Notes in Comput. Sci., Springer, 2000, pp. 176–187.
[25] A. Razborov, The gap between the chromatic number of a graph and the rank of its adjacency matrix is superlinear, Discrete Math. 108 (1992) 393–396.
[26] M. Sajna, Cycle decompositions of K_n and $K_n - I$, Ph.D. Thesis, Simon Fraser University, Canada, 1999.
[27] M. Tait, The Alon–Saks–Seymour and Rank-Coloring Conjectures, M.Sc. Thesis, University of Delaware, 2011, p. 57.
[28] H. Tverberg, On the decomposition of K_n into complete bipartite graphs, J. Graph Theory 6 (1982) 493–494.
[29] J.H. van Lint, $\{0, 1, *\}$ distance problems in combinatorics, in: Surveys in Combinatorics 1985 (Glasgow, 1985), in: London Math. Soc. Lecture Note Ser., vol. 103, Cambridge Univ. Press, Cambridge, 1985, pp. 113–135.
[30] J.H. van Lint, R.M. Wilson, A Course in Combinatorics, second ed., Cambridge University Press, Cambridge, 2001, p. xiv+602.
[31] C. van Nuffelen, Research problems: a bound for the chromatic number of a graph, Amer. Math. Monthly 83 (4) (1976) 265–266.
[32] S. Vishwanathan, A polynomial space proof of the Graham–Pollak theorem, J. Combin. Theory Ser. A 115 (2008) 674–676.
[33] S. Vishwanathan, A counting proof of the Graham Pollak theorem, Preprint. Available online at: http://arxiv.org/abs/1007.1553.
[34] D.B. West, Introduction to Graph Theory, Prentice Hall, 1996, p. xvi+512.