A Turán Type Problem Concerning the Powers of the Degrees of a Graph

Yair Caro *
and
Raphael Yuster †
Department of Mathematics
University of Haifa-ORANIM, Tivon 36006, Israel.

AMS Subject Classification: 05C35,05C07 (primary).

Submitted: April 18, 2000; Accepted: August 30, 2000

Abstract

For a graph G whose degree sequence is $d_1, \ldots, d_n$, and for a positive integer p, let $e_p(G) = \sum_{i=1}^n d_i^p$. For a fixed graph H, let $t_p(n,H)$ denote the maximum value of $e_p(G)$ taken over all graphs with n vertices that do not contain H as a subgraph. Clearly, $t_1(n,H)$ is twice the Turán number of H. In this paper we consider the case $p > 1$. For some graphs H we obtain exact results, for some others we can obtain asymptotically tight upper and lower bounds, and many interesting cases remain open.

1 Introduction

All graphs considered here are finite, undirected, and have no loops or multiple edges. For the standard graph-theoretic notations the reader is referred to [1]. For a graph G whose degree sequence is $d_1, \ldots, d_n$ let $e_p(G) = \sum_{i=1}^n d_i^p$. Clearly, $e_1(G) = 2e(G)$. Recently, several papers were published concerning the problem of maximizing $e_2(G)$ over all graphs having n vertices and m edges. See, e.g., [2, 3, 9, 4, 10]. In this line of research no restriction is imposed on the structure of G. Along the spirit of Turán Theory we consider the problem of finding the maximum of $e_p(G)$ over the class of graphs which contain no copy of prescribed forbidden

*e-mail: yairc@macam98.ac.il
†e-mail: raphy@macam98.ac.il

subgraphs. For a fixed graph H, let $t_p(n,H)$ denote the maximum value of $e_p(G)$ taken over all graphs with n vertices that do not contain H as a subgraph. Clearly, $t_1(n,H) = 2t(n,H)$ where $t(n,H)$ is the Turán Number of H.

The Turán number $t(n,H)$ is one of the most studied parameters in Graph Theory. Many interesting and non-trivial results give either exact values or asymptotically tight upper and lower bounds for $t(n,H)$. For example, the classic result of Turán (from which Turán Theory has emerged) determines $t(n,K_p)$ for all n and p. There are still many open problems, even when H is a rather simple graph. For example, when H is a tree with k vertices it is conjectured that $t(n,H) = (k/2-1)n(1+o(1))$, and when $k-1$ divides n the conjecture is that $t(n,H) = (k/2-1)n$. The lower bound is obtained by taking $n/(k-1)$ vertex-disjoint copies of K_{k-1}. The upper bound would follow if one can prove the famous conjecture of Erdős and Sós [5], which states that graphs with $(k/2-1)n+1$ edges contain every tree with k vertices. This conjecture is known to hold if G is C_4-free [11] or if the tree has a vertex adjacent to at least $(k-2)/2$ leaves [12].

In many cases, the extremal graphs with respect to $t(n,H)$ tend to be regular or almost regular. That is, the k'th central moment of the degree sequence is either zero or very small. If we wish to investigate highly non-regular H-forbidden graphs, then just counting the number of edges does not suffice. If we wish to maximize the second central moment of the degree sequence of H-forbidden graphs, then the parameter $t_2(n,H)$ is the correct measure. Likewise, for the p'th central moment the parameter $t_p(n,H)$ is the suitable one. In this paper we consider $t_p(n,H)$ for $p > 1$. For some graphs H we are able to give exact or near-exact results, while for others the problem remains open.

Our first result shows that when $H = K_k$, the extremal graph that yields $t_p(n,H)$ is exactly the same graph that yields $t_1(n,H)$, namely the Turán Graph $T(n,k)$.

Theorem 1.1 *Let $k > 2$ be a positive integer, and let $p \geq 1$. Then, $t_p(n,K_k) = e_p(T(n,k))$, where $T(n,k)$ is the Turán Graph.*

Theorem 1.1 shows that the parameter p plays no role in the extremal graphs for $t_p(n,K_k)$. This is no longer true when we consider paths. Let P_k denote the path with k vertices. Faudree and Schelp [7] characterized the extremal graphs that yield $t(n,P_k)$. Let $r \equiv n \bmod (k-1)$. An extremal graph giving $t(n,P_k)$ is obtained by taking $\lfloor n/(k-1)\rfloor$ vertex-disjoint copies of cliques of order $k-1$ and, if $r \neq 0$, another clique K_r on the remaining vertices. Hence, $t(n,P_k) = \binom{k-1}{2}\lfloor n/(k-1)\rfloor + \binom{r}{2}$. These graphs are far from optimal when considering $t_p(n,P_k)$ when $p > 1$, since they have small maximum degree.

Our next theorem determines $t_p(n,P_k)$ for n sufficiently large (for small values of n there are some disturbances). In order to describe this theorem we define the graph $H(n,k)$ for $n \geq k \geq 4$ as follows. The vertex set of H is composed of two parts A and B where $|B| = \lfloor k/2\rfloor - 1$ and $|A| = n - |B|$. B induces a complete graph, and A induces an independent set when k is even, or a single edge plus $|A|-2$ isolated vertices when k is odd. All possible edges between A and B exist.

Theorem 1.2 *Let $k \geq 4$, let $p \geq 2$ and let $n > n_0(k)$. Then, $H(n,k)$ contains no copy of P_k and $t_p(n, P_k) = e_p(H(n,k))$. Furthermore, $H(n,k)$ is the unique extremal graph.*

Note that, trivially, $t_p(n, P_2) = 0$ and $t_p(n, P_3) = n$ when n is even and $t_p(n, P_3) = n - 1$ when n is odd (by taking a maximum matching on n vertices). Also note that when n is small compared to k, the graph $H(n,k)$ is *not* the extremal graph. As an extreme example note that $t_p(k-1, P_k) = (k-1)(k-2)^p$ and is obtained by K_{k-1}. A close examination of the proof of Theorem 1.2 shows that the value of $n_0(k)$ in the statement of the theorem is $O(k^2)$. Another thing to note is that, as long as $p \geq 2$, the actual value of p is insignificant.

Let C^* be the family of even cycles. It is an easy exercise to show that any graph with more than $\lfloor 3(n-1)/2 \rfloor$ edges contains an even cycle. This bound is sharp and there are exponentially many extremal graphs [1]. In fact, the extremal graphs can be constructed recursively as follows. For $n = 1$ take a single point. For $n = 2$ take a single edge. If $n > 2$ we construct graphs with no even cycles and with $\lfloor 3(n-1)/2 \rfloor$ edges as follows. Let G be any such extremal graph with $n-2$ vertices. Pick any vertex x of G and add to G two new vertices a, b. Now add a triangle on x, a, b. The resulting graph has n vertices $e(G)+3 = \lfloor 3(n-1)/2 \rfloor$ edges, and no even cycle. Notice that the *Friendship Graph* F_n is one of the extremal graphs. F_n is defined as follows. Take a star with n vertices and add a maximum matching on the set of leaves. Thus, F_n has exactly $n - 1 + \lfloor (n-1)/2 \rfloor = \lfloor 3(n-1)/2 \rfloor$ edges, and no even cycle. Note that when n is odd, $e_2(F_n) = (n-1)^2 + 4(n-1) = n^2 + 2n - 3$ and when n is even $e_2(F_n) = (n-1)^2 + 4(n-2) + 1 = n^2 + 2n - 6$. Our next theorem shows that, unlike the Turán case, there is only one extremal graph giving $t_2(n, C^*)$, and it is F_n. (Notice the natural extension of the definition of t_p to families of graphs).

Theorem 1.3 *For n sufficiently large, $t_2(n, C^*) = e_2(F_n)$ and F_n is the unique extremal graph.*

We mention that Theorem 1.3 also holds for $p > 2$, but the proof is rather technical and we omit it.

The rest of this paper is organized as follows. In Section 2 we consider complete graphs and prove Theorem 1.1. In Section 3 we consider paths and prove Theorem 1.2. Some other acyclic graphs H for which $t_p(n, H)$ can be determined are handled in Section 4. In Section 5 we prove Theorem 1.3 and also asymptotically determine $t_k(n, K_{k,k})$. The final section contains some concluding remarks and open problems.

2 Complete graphs

In order to prove Theorem 1.1 we need the following theorem of Erdős [6] that characterizes the maximal degree sequences of graphs without a K_k.

Lemma 2.1 (Erdős [6]) *Let $G = (V, E)$ be a graph without a K_k. Then, there is a $(k-1)$-partite graph $G' = (V, E')$ such that for every $v \in V$, $d_G(v) \leq d_{G'}(v)$.* $\Box$

If G and G' are as in Lemma 2.1 then, clearly, $e_p(G) \le e_p(G')$ for all $p \ge 1$. Thus, the following corollary immediately follows from Lemma 2.1:

Corollary 2.2 *For every $n \ge k-1 \ge 1$ there exists a complete $(k-1)$-partite graph G with n vertices such that $t_p(n, K_k) = e_p(G)$.* □

Proof of Theorem 1.1: It suffices to show that the complete $(k-1)$-partite graph G in Corollary 2.2 is the Turán Graph $T(n, k)$. For $k = 2$ this is trivial. Assume therefore that $k \ge 3$. It suffices to show that if G' is any complete $(k-1)$-partite graph that has (at least) two vertex classes X and Y with $|X| - |Y| > 1$ then the complete $(k-1)$-partite graph G'' obtained from G' by transferring a vertex from X to Y has $e_p(G'') > e_p(G')$. Indeed, putting $|X| = x$ and $|Y| = y$ we have

$$e_p(G'') - e_p(G') = (y+1)(n-y-1)^p + (x-1)(n-x+1)^p - y(n-y)^p - x(n-x)^p > 0.$$

where the last inequality may be verified using standard (although tedious) calculus, and the facts that $n \ge x + y$ and $x - y - 1 > 0$. For example, if $p = 2$ the expression in the middle of the last inequality is equivalent to the expression $(x-y-1)(n+3(n-x-y))$. □

Let K'_k be the graph obtained from K_k by adding a new vertex of degree one, connected to one of the original vertices. It is not difficult to show that $t_p(n, K'_k) = t_p(n, K_k)$ for $n > n(k)$ (assuming $k \ge 3$ and $p \ge 2$). Indeed, we can state this more generally.

Proposition 2.3 *Let H be a vertex-transitive graph with at least two edges. Let H' be obtained from H by adding a new vertex of degree one connected to one of the original vertices. Then, if $p \ge 2$, $t_p(n, H') = t_p(n, H)$ for $n > n(p, H)$.*

Proof: Clearly, $t_p(n, H') \ge t_p(n, H)$ since H' contains H. Now assume that equality does not hold. Let G be an n-vertex graph having a copy of H as a subgraph, but having no H' as a subgraph, and having $e_p(G) = t_p(n, H') > t_p(n, H)$. Since H is vertex-transitive, the set of vertices of every copy of H in G is disconnected from the other vertices of G, since otherwise we would have an H'. Thus, if t is the number of copies of H in G we have $e_p(G) \le th(h-1)^p + t_p(n - ht, H)$. However, $t_p(n, H) = \Omega(n^p)$ as can be seen by the star S_n which has no copy of H (recall that H is vertex-transitive with at least two edges), and $e_p(S_n) = \Omega(n^p)$, thus for n sufficiently large, $th(h-1)^p + t_p(n - ht, H)$ is maximized when $t = 0$. Consequently, $e_p(G) \le t_p(n, H)$, a contradiction. □

3 Paths

In order to determine $t_p(n, P_k)$ it is useful to have an upper bound on the maximum number of edges possible in a graph not containing P_k. The following lemma, which is a theorem of Faudree and Schelp, determines the Turán number for paths.

Lemma 3.1 (Faudree and Schelp [7]) *Let $k > 1$ and let $n > 0$. Let $r = \lfloor n/(k-1) \rfloor$ and let $s \equiv n \bmod (k-1)$ where $0 \leq s < k-1$. Then $t(n, P_k) = r\binom{k-1}{2} + \binom{s}{2}$.* □

In fact, Faudree and Schelp also characterized the extremal Turán graphs. The graph composed of r vertex-disjoint cliques of order $k-1$ plus an additional clique of order s is extremal (sometimes, however, it is not the only extremal graph). It would be somewhat more convenient to use the following less accurate upper bound for $t(n, P_k)$, that is always at least as large as the value in Lemma 3.1

Corollary 3.2 *If G has n vertices and is P_k-free then $e(G) \leq n(k-2)/2$.* □

We also need a lemma bounding $e_p(G)$ for n-vertex graphs G that have linearly many edges, and have maximum degree $\Theta(n)$.

Lemma 3.3 *Let $p \geq 2$ be an integer, let $0.5 < \alpha \leq 1$ and let $t > \alpha$ be real. Let G be an n-vertex graph with $\Delta(G) \leq \alpha n$ and with at most tn edges. Then:*

$$e_p(G) \leq \frac{t}{\alpha}(\alpha n)^p + o(n^p).$$

Proof: Consider the degree sequence of G, denoted $\{d_1, \ldots, d_n\}$. It is a sequence of n nonnegative integers whose sum is at most $2tn$ and whose elements do not exceed αn. If we "forget" that this sequence is graphic then $d_1^p + \ldots + d_n^p$ is, obviously, at most

$$\frac{2tn}{\alpha n}(\alpha n)^p = \frac{2t}{\alpha}(\alpha n)^p.$$

However, the sequence is graphic. This means, for example, that if there is a vertex of degree, say, αn, then there are at least αn nonnegative elements in the sequence. In fact, for any fixed $\beta \leq \alpha$ there are at most $t/\beta + o(1)$ vertices with degree at least βn, and if this happens, then the other degrees are all at most $t/\beta + o(1)$ which is constant (and hence have no significant contribution to $e_p(G)$). By the convexity of the polynomial x^p, the optimal situation is obtained by taking vertices with degree αn as many as possible (there are at most $t/\alpha + o(1)$ such vertices), and this forces the other vertices (except maybe one) to have constant degree. Hence,

$$e_p(G) \leq \frac{t}{\alpha}(\alpha n)^p + o(n^p).$$

□

Before we prove Theorem 1.1 we need to dispose of the special case $k = 5$, since this value causes technical difficulties in the proof.

Lemma 3.4 *For $n \geq 12$, $t_p(n, P_5) = e_p(H(n, 5))$. Furthermore, $H(n, 5)$ is the unique extremal graph.*

Proof: First note that for $r \geq 3$, $H(r,5)$ is a star with r vertices with an additional edge connecting two of its leaves. Assume G is an n-vertex graph having no P_5 and $t_p(n,P_5) = e_p(G)$. Trivially, G may only contain cycles of length 3 or 4. In fact, the vertices of every 4-cycle must induce a component of G isomorphic to K_4, and for any 3-cycle, the connected component to which it belongs must be an $H(r,5)$ for some $r \geq 3$. Thus, the components of G are either K_4's or $H(r,5)$'s (there may be several with distinct values of r), or trees. Trivially, every tree T with r vertices has $e_p(T) \leq e_p(S_r)$ where S_r is the r-vertex star. Similarly, $e_p(S_r) \leq e_p(H(r,5))$ since $H(r,5)$ contains S_r (if $r=1$ or $r=2$ we define $H(1,5) = S_1$ and $H(2,5) = S_2$). Thus, we may assume that every component is either a K_4 or an $H(r,5)$. Another trivial check is that $e_p(H(r_1,5)) + e_p(H(r_2,5)) < e_p(H(r_1+r_2,5))$. Thus, we can assume that there is at most one component equal to $H(r,5)$ and the other components are K_4. In fact, replacing three copies of K_4 (contributing $12 \cdot 3^p$ to $e_p(G)$ with one copy of $H(12,5)$ (contributing $11^p + 2^{p+1} + 9$ to $e_p(G)$) improves $e_p(G)$ so we can assume that there are at most two components isomorphic to K_4. Since $n \geq 12$ we must have $r \geq 4$. Now, for $r \geq 4$, $e_p(H(r+4,5)) - e_p(H(r,5)) > 4 \cdot 3^p$ so it is better to replace an $H(r,5)$ and a K_4 with one $H(r+4,5)$. Consequently, $G = H(n,5)$. $\square$

Proof of Theorem 1.2: In the proof we shall assume, wherever necessary, that n is sufficiently large as a function of k, and that $k \neq 5$. It is trivial to check that the graph $H(n,k)$ defined in the introduction has no P_k. We therefore have the lower bound $t_p(n,P_k) \geq e_p(H(n,k))$. To prove the theorem it suffices to show that any P_k-free graph G with n vertices that is not $H(n,k)$ has $e_p(G) < e_p(H(n,k))$ for every $p \geq 2$. Assume the contrary, and let $G = (V,E)$ be a P_k-free graph with n vertices that is maximal in the sense that $e_p(G) = t_p(n,P_k)$ and $G \neq H(n,k)$. We will show how to derive a contradiction.

According to Corollary 3.2, $|E| \leq n(k-2)/2$. Order the vertices of G in nonincreasing degree order. That is $V = \{x_1,\ldots,x_n\}$ where $d_G(x_i) \geq d_G(x_{i+1})$ for $i = 1,\ldots,n-1$. Put $d_i = d_G(x_i)$, and put $b = \lfloor k/2 \rfloor - 1$. Note that

$$e_p(H(n,k)) = bn^p + o(n^p) \tag{1}$$

Put $B = \{x_1,\ldots,x_b\}$ and $A = \{x_{b+1},\ldots,x_n\}$. First observe that we may assume that d_1 is very large. For instance, we may assume that for all $k \neq 5$, $d_1 > 0.79n$ since otherwise, applying Lemma 3.3 to G with $\alpha = 0.79$ and $t = (k-2)/2$ we get for $k \neq 5$,

$$e_p(G) \leq \frac{k/2-1}{0.79}(0.79)^p n^p + o(n^p) \leq (0.395k - 0.79)n^p + o(n^p) < bn^p + o(n^p) = e_p(H(n,k)).$$

Lemma 3.5 *If $d_b \leq 0.65n$ then $e_p(G) < e_p(H(n,k))$.*

proof: By (1) it suffices to show that $e_p(G) \leq cn^p + O(n^{p-1})$ where c is a constant smaller than b. Consider the spanning subgraph of G obtained by deleting all the edges adjacent with the vertices of $B \setminus \{x_b\}$. Denote this subgraph by G'. The maximum degree of G' is at most $0.65n$. Let f_i denote the degree of x_i in G' for $i = b,\ldots,n$. By definition,

$e_p(G') = f_b^p + \ldots + f_n^p$. Since $f_i \geq d_i - b + 1$, and since $f_b + \ldots + f_n \leq n(k-2) = O(n)$ we have

$$e_p(G) = d_1^p + \ldots + d_n^p \leq d_1^p + \ldots + d_{b-1}^p + (f_b + b - 1)^p + \ldots + (f_n + b - 1)^p = \qquad (2)$$
$$d_1^p + \ldots + d_{b-1}^p + e_p(G') + o(n^p)$$

Define $t = e(G')/n$. We consider three cases according to the value of t.

Case 1: $t < 0.65$. Since the degree sequence has sum at most $1.3n$ and no element is larger than $0.65n$ we have $e_p(G') \leq 2(0.65n)^p$ and using (2) we get:

$$e_p(G) \leq d_1^p + \ldots + d_{b-1}^p + e_p(G') + o(n^p) \leq (b - 1 + 2(0.65)^p)n^p + o(n^p) < e_p(H(n,k)).$$

Case 2: $1.45 > t \geq 0.65$. According to Lemma 3.3 with $\alpha = 0.65$ we know that:

$$e_p(G') \leq \frac{t}{0.65}(0.65n)^p + o(n^p) \leq 0.9425n^p + o(n^p).$$

Using (2) we get:

$$e_p(G) \leq d_1^p + \ldots + d_{b-1}^p + e_p(G') + o(n^p) \leq (b - 1 + 0.9425)n^p + o(n^p) < e_p(H(n,k)).$$

Case 3: $t \geq 1.45$. According to Lemma 3.3 with $\alpha = 0.65$ we know that:

$$e_p(G') \leq \frac{t}{0.65}(0.65n)^p + o(n^p). \qquad (3)$$

Let z denote the number of edges of G with both endpoints in $B \setminus \{x_b\}$. Clearly, $z \leq \binom{b-1}{2} < k^2$. Now,

$$n\frac{k-2}{2} \geq e(G) = e(G') + d_1 + \ldots + d_{b-1} - z = tn + d_1 + \ldots + d_{b-1} - z \geq tn + d_1 + \ldots + d_{b-1} - k^2$$

It follows that

$$d_1 + \ldots + d_{b-1} \leq n(\frac{k-2}{2} - t) + k^2 = n(\frac{k-2}{2} - t) + o(n).$$

Since $d_i < n$ the last inequality immediately gives:

$$d_1^p + \ldots + d_{b-1}^p \leq (\frac{k-2}{2} - t)n^p + o(n^p). \qquad (4)$$

Plugging (3) and (4) in (2) yields:

$$e_p(G) \leq d_1^p + \ldots + d_{b-1}^p + e_p(G') + o(n^p) \leq$$

$$(\frac{k-2}{2} - t + \frac{t}{0.65}(0.65)^p)n^p + o(n^p) \leq (\frac{k-2}{2} - 0.5075)n^p + o(n^p) < e_p(H(n,k)).$$

□

In view of Lemma 3.5 we may now assume $d_b > 0.65n$, and due to the remark prior to Lemma 3.5 we may also assume that when $k \neq 5$, $d_1 > 0.79n$. Let $A' \subset A$ be the set of vertices that have a neighbor in B. Let $G[A']$ denote the subgraph induced by A'.

Lemma 3.6 *If k is even, then $G[A']$ has no edges. If k is odd, then $G[A']$ contains at most one edge.*

Proof: Assume the contrary. We will derive a contradiction by showing that G contains a P_k. We distinguish three cases
Consider first the case where k is even. Let (a_0, a_1) be an edge of $G[A']$. By the definition of A', a_1 has a neighbor in B. Assume w.l.o.g. x_1 is a neighbor of a_1. Note that since $d_i > 0.65n$ for $i = 1, \dots, b$ we have that any two vertices of B have at least $0.3n$ common neighbors in G, and hence at least $0.3n - (b-2) > k$ common neighbors in A'. Therefore, let $a_i \in A'$ be a common neighbor of x_{i-1} and x_i for $i = 2, \dots, b$ such that $a_0, a_1, \dots, a_b$ are all distinct. Let $a_{b+1} \in A'$ be a neighbor of x_b distinct from $a_0, \dots, a_b$. We have that $a_0, a_1, x_1, a_2, x_2, a_3, \dots, a_{b-1}, x_{b-1}, a_b, x_b, a_{b+1}$ is a P_k.
Next, consider the case where k is odd and there are two edges in $G[A']$ sharing a common endpoint in A'. Denote these two edges by (a_{-1}, a_0) and (a_0, a_1). As in the previous case we can obtain a P_k of the form $a_{-1}, a_0, a_1, x_1, a_2, x_2, a_3, \dots, a_{b-1}, x_{b-1}, a_b, x_b, a_{b+1}$.
Next, consider the case where k is odd and $G[A']$ has two independent edges, denoted (a_0, a_1) and (a_{b+1}, a_{b+2}) such that a_1 and a_{b+1} have at least two vertices of B in their neighborhood union. W.l.o.g. a_1 is a neighbor of x_1 and a_{b+1} is a neighbor of x_b. As in the previous cases we can obtain a P_k of the form $a_0, a_1, x_1, a_2, x_2, a_3, \dots, a_{b-1}, x_{b-1}, a_b, x_b, a_{b+1}, a_{b+2}$.
The only remaining case is that k is odd and $G[A']$ contains two or more independent edges, and all the endpoints of these independent edges are connected to a single vertex of B, say, x_1. In this case, there may not be a P_k present, but we will show that there is a P_k-free graph G' on n vertices with $e_p(G') > e_p(G)$, contradicting the maximality of G. Since we assume $k \geq 7$ we have $b \geq 2$ so $x_2 \in B$. Let A^* denote the set of non-isolated vertices in $G[A']$. $|A^*| \geq 4$ and no vertex of A^* is connected to x_2. We may delete the $|A^*|/2$ independent edges of $G[A']$, and replace them with $|A^*|$ new edges from x_2 to each of the vertices of A^*. Clearly, if G is P_k-free, so is G' (this follows from the fact that k is odd, so $b = \lfloor k/2 \rfloor - 1 = (k-3)/2$). However, the degree sequence of G' majorizes that of G since the degree of x_2 increased, while the other degrees have not changed. Hence, $e_p(G') > e_p(G)$, a contradiction. □

An immediate corollary of Lemma 3.6 is the following:

Corollary 3.7 *The subgraph of G induced by $B \cup A'$ is a spanning subgraph of $H(b+a', k)$ where $|A'| = a'$. In particular, if $A' = A$ then G is a spanning subgraph of $H(n, k)$.* □

Note that by Corollary 3.7 we have that if $A' = A$ then $e_p(G) < e_p(H(n,k))$ since $G \neq H(n,k)$. This contradicts the maximality of $e_p(G)$. The only remaining case to consider is when $A' \neq A$. The following lemma shows that this is impossible, due to the maximality of G. This final contradiction completes the proof of Theorem 1.2.

Lemma 3.8 *If $A' \neq A$ then there exists a P_k-free graph G' with n vertices such that $e_p(G) < e_p(G')$.*

Put $A'' = A \setminus A'$. We claim that each $v \in A''$ has at most one neighbor in A'. Indeed, if it had two neighbors, say a_0, a_1 then, as in the previous cases, we can obtain a P_k of the form $a_0, v, a_1, x_1, a_2, x_2, \ldots, a_{b-1}, x_{b-1}, a_b, x_b, a_{b+1}$ (in fact, if k is even this is a P_{k+1}). Since $G[A'']$ is P_k-free it contains at most $(k/2-1)a''$ edges, where $a'' = |A''| = a - a'$. Hence, it contains a vertex v whose degree is at most $k-2$. Hence $d_G(v) \le k-1$. Delete all edges adjacent to v in G, and connect v to all edges of B. Denote the new graph by G'. Note that G' is also P_k-free. To see this, note that otherwise, any P_k in G' must contain v. Let $x_i \in B$ be a neighbor of v in such a P_k. If v is not an endpoint of the P_k it also contains another neighbor $x_j \in B$ in the path. Since x_i and x_j have many common neighbors in A' (much more than k), let $v' \in A'$ be such a common neighbor which is not on the P_k (if v is an endpoint of the P_k it suffices to take $v' \in A'$ to be any neighbor of x_i not on the P_k). Replacing v with v' on the P_k we obtain a P_k in G, contradicting the assumption. We now show that $e_p(G') > e_p(G)$. Consider the effect of the transformation from G to G' on the degree sequence. The degrees of the vertices of B increased by one. The degree of v may have decreased by at most $k-1-b$. The degrees of the neighbors of v in G have decreased by 1. Since every vertex of B has degree at least $0.65n$, the total increase in $e_p(G') - e_p(G)$ contributed by the vertices of B is at least

$$b((0.65n+1)^p - (0.65n)^p) = bp(0.65n)^{p-1} + o(n^{p-1}).$$

Assuming $k \neq 5$, we know $d_1 > 0.79n$. This implies that $a'' < 0.21n$. This fact, together with Lemma 3.6 shows that every vertex of A'' has degree at most $0.21n$ in G. Thus, the total decrease in $e_p(G') - e_p(G)$ contributed by the vertices of A is at most

$$(k-1)((0.21n)^p - (0.21n-1)^p) + (k-1)^2 - b^2 = (k-1)p(0.21n)^{p-1} + o(n^{p-1})$$

Hence, for $k \neq 5$

$$e_p(G') - e_p(G) \ge p(b(3.09)^{p-1} - k + 1)(0.21)^{p-1}n^{p-1} + o(n^{p-1}) > 0. \qquad \square$$

4 Other acyclic graphs

A linear forest is a forest whose components are paths. An even linear forest is a forest whose components are paths with an even number of vertices (distinct components may have different lengths). The simplest example of an even linear forest is a *matching*, namely, a graph whose components are single edges. Let M_k denote the matching with $2k$ vertices. Note that every even linear forest F with $2k$ vertices is a spanning subgraph of P_{2k} and contains M_k as a spanning subgraph. Thus, for every n we have $t_p(n, M_k) \le t_p(n, F) \le t_p(n, P_{2k})$. We immediately get the following proposition:

Proposition 4.1 *Let $k \ge 2$ be an integer, and let $p \ge 2$ be an integer. If F is an even linear forest with $2k$ vertices then, for n sufficiently large, $t_p(n,F) = e_p(H(n,2k))$ where $H(n,2k)$ is the extremal graph appearing in Theorem 1.2.*

Proof: By Theorem 1.2 we know that for n sufficiently large, $t_p(n, P_{2k}) = e_p(H(n, 2k))$. On the other hand, it is trivial to check that $H(n, 2k)$ does not contain M_k as a subgraph. Hence, $t_p(n, M_k) \geq e_p(H(n, 2k))$. Since $t_p(n, M_k) \leq t_p(n, F) \leq t_p(n, P_{2k})$ we must have $t_p(n, M_k) = t_p(n, F) = t_p(n, P_{2k})$ for n sufficiently large. $\Box$

Another family of trees for which t_p is easy to compute is the family of stars. Indeed, let S_k denote the star with $k \geq 2$ vertices. Clearly, if G has no S_k it has $\Delta(G) \leq k-2$. Thus, every n-vertex graph G that is $k-2$-regular must satisfy $t_p(n, S_k) = e_p(G)$. If $n > k-2$ is even then it is well-known that such G exist for all $k \geq 2$. (in fact, they can be obtained by an edge-disjoint union of $k-2$ perfect matchings). So is the case when n is odd and k is even (they can be obtained by an edge-disjoint union of $(k-2)/2$ Hamilton cycles). If both n and k are odd then there do not exist $k-2$-regular n-vertex graphs, so, clearly, if G has $n-1$ vertices with degree $k-2$ and one vertex with degree $k-3$, then $t_p(n, S_k) = e_p(G)$. Such G are well-known to exist for all $n > k-2$. In fact, they can be obtained by an edge-disjoint union of $(k-3)/2$ Hamilton cycles plus a maximum matching. Note that if $n \leq k-2$, then, clearly, $t_p(n, S_k) = e_p(K_n)$. To summarize:

Proposition 4.2 *Let S_k be the star with $k \geq 2$ vertices. Then:*

1. *If $n \leq k-2$ then $t_p(n, S_k) = n(n-1)^p$.*
2. *If $n > k-2$ and nk is even then $t_p(n, S_k) = n(k-2)^p$.*
3. *If $n > k-2$ and nk is odd then $t_p(n, S_k) = (n-1)(k-2)^p + (k-3)^p$.* $\Box$

A slight modification of S_k is the *near star* S_k^*. This graph is an S_{k-1} to which we add one new neighbor to one of the leaves. So, e.g., $S_4^* = P_4$. This slight modification to S_k yields an entirely different result for $t_p(n, S_k^*)$.

Proposition 4.3 *If $n > 2k$ then $t_p(n, S_k^*) = e_p(S_n) = (n-1)^p + (n-1)$.*

Proof: Let G be a graph without an S_k^*. If G has a vertex of degree $m \geq k-1$ then, trivially, this vertex belongs to a component of G that is an S_{m+1}, since otherwise G would have an S_k^*. Hence, each component of G either has maximum degree at most $k-2$, or else is a star. Let s denote the number of vertices of G that belong to components of the first type. Then, $e_p(G) \leq s(k-2)^p + (n-s-1)^p + (n-s-1)$. Clearly, when $n > 2k$ (in fact, even before that point as p increases), the last inequality is optimized when $s = 0$. Thus, $e_p(G) \leq (n-1)^p + (n-1)$. Equality is obtained since S_n is S_k^*-free. $\Box$

A connected bipartite graph is *equipartite* if the two vertex classes forming the bipartition have equal size. For equipartite trees T that obey the Erdős-Sós Conjecture we can asymptotically determine $t_p(n, H)$. Examples of such trees are even paths (however, for these we already have the sharp result of Theorem 1.2), but there are many others. One example is the balanced double star $S_{k,k}$, that is obtained by taking two disjoint copies of the star S_k and joining their roots with an edge. Sidorenko [12] has proved that the Turán number of $S_{k,k}$ satisfies $t(n, S_{k,k}) \leq (k-1)n$ (equality is obtained when $2k-1$ divides n). Namely, the Erdős-Sós Conjecture holds for $S_{k,k}$.

Proposition 4.4 *If H is an equipartite tree with $2k$ vertices, and $t(n,H) \le (k-1)n$ then*

$$t_p(n,H) = (k-1)n^p + o(n^p).$$

Proof: We use Lemma 3.3 with $\alpha = 1$ and $t = k-1$. Indeed, if G is an n-vertex graph that is H-free, then G has at most tn edges. Thus, by Lemma 3.3, $e_p(G) \le (k-1)n^p + o(n^p)$. Consequently, $t_p(n,H) \le (k-1)n^p + o(n^p)$. On the other hand, consider the complete bipartite graph $B_{k-1,n-k+1}$. Since H is equipartite, $B_{k-1,n-k+1}$ does not contain H as a subgraph. Since $e_p(B_{k-1,n-k+1}) = (k-1)(n-k+1)^p + (n-k+1)(k-1)^p = (k-1)n^p + o(n^p)$ we have $t_p(n,H) = (k-1)n^p + o(n^p)$. $\square$

5 Even cycles and complete bipartite graphs

Proof of Theorem 1.3 Let G have n vertices and no even cycle, and assume that $G \ne F_n$. We must show $e_2(G) < e_2(F_n)$. Let $d_1 \ge d_2 \ge \ldots \ge d_n$ be the degree sequence of G, and let $x_1, \ldots, x_n$ be the corresponding vertices. Notice first that $d_1 + d_2 \le n+1$. Indeed, otherwise x_1 and x_2 would have two distinct common neighbors, and G would contain a C_4.
We first consider the case $d_1 \le 0.75n$. Straightforward convexity arguments, plus the fact that $d_1 + d_2 \le n+1$ and the fact that $e(G) < 1.5n$ show that the largest possible value for the sum of squares is at most the one given by a sequence of the form:

$$0.75n\ ,\ 0.25n(1+o(1))\ ,\ 0.25n(1+o(1))\ ,\ 0.25n(1+o(1)\ ,\ d_5\ ,\ d_6\ ,\ \ldots\ ,\ d_n$$

where d_i is bounded by the constant 4 for $i \ge 5$. Hence $e_2(G) \le \frac{12}{16}n^2 + o(n^2) < e_2(F_n)$ for n sufficiently large (in fact, $n = 12$ already suffices).
Next, we consider the case $d_1 > 0.75n$. Consider any nonincreasing sequence of n nonnegative integers having the following properties:

1. $0.75n < d_1 \le n-1$.
2. $d_1 + d_2 \le n+1$.
3. There are at least $d_1 + 1$ nonzero elements in the sequence.
4. The sum of the elements is at most $3(n-1)$.

Putting $d_1 = x$, the degree sequence dominates the sequence $S = \{x, 1, 1, 1, \ldots, 1, 0, 0, \ldots, 0\}$ (there are x ones and $n-x-1$ zeroes here). Hence, there are at most $3(n-1) - 2x$ units to assign to S (subject to the four properties above) in order to obtain the degree sequence of G. By convexity, the sum of squares is maximized if we assign $n-x$ additional units to the second, third, etc. elements of S, until we run out of units. Thus,

$$e_2(G) \le x^2 + (n+1-x)^2 \frac{3(n-1)-2x}{n-x} + 1^2\left(x - \frac{3(n-1)-2x}{n-x}\right).$$

Putting $x = n - k$ (where $1 \le k < n/4$) the r.h.s. of the last inequality is equal to $n^2 - (k-3)n + 3k^2 - 6$. Thus,

$$e_2(G) \le n^2 - (k-3)n + 3k^2 - 6.$$

Note that when $n > \max\{20\ ,\ 4k\}$ and $k \ge 2$ we have $n^2 - (k-3)n + 3k^2 - 6 < n^2 + 2n - 6 \le e_2(F_n)$. Thus, we have shown that if $d_1 < n - 1$ then $e(G) < e_2(F_n)$. If $d_1 = n - 1$ then, subject to the above four properties, the sum of squares is maximized by the unique sequence $n-1, 2, 2, \ldots, 2$. When n is odd there is only one graph with this degree sequence, namely F_n, and we assume $G \ne F_n$ so $e_2(G) < e_2(F_n)$. When n is even this is not the degree sequence of any graph (as the sum of the elements is odd), thus, subject to the above four properties and the requirement that the sequence be graphic, the sum of squares is maximized by the sequence $n-1, 2, 2, \ldots, 2, 1$. There is only one graph with this degree sequence, namely F_n. Again by our assumption, $G \ne F_n$, so $e_2(G) < e_2(F_n)$. $\Box$

We now turn our attention to complete bipartite graphs. The Turán number for $K_{k,k}$ is well-understood only for $k = 2$. It is known that $t(K_{2,2}, n) = 0.5n^{3/2}(1 + o(1))$ (cf. [1]). Exact values and extremal graphs are known only in special cases. Recently, Füredi proved in [8] that if a graph has $q^2 + q + 1$ vertices $q > 13$, m edges and no $K_{2,2}$ then $m \le 0.5q(q+1)^2$, and equality holds for graphs obtained from finite projective planes with polarities. If $k \ge 3$ the asymptotic behavior of $K_{k,k}$ is not known. The best (and rather simple) bounds are $t(K_{k,k}, n) \le O(n^{2-1/k})$ and $t(K_{k,k}, n) \ge \Omega(n^{2-2/k})$. Our next proposition shows that $t_k(n, K_{k,k})$ can be asymptotically determined for every $k \ge 2$. In fact, something slightly stronger can be proved:

Proposition 5.1 *Let $2 \le a \le k$ where a, k are integers. Then: $t_k(n, K_{a,k}) = (a-1)n^k(1 + o(1))$. Furthermore, if $p \ge k$ then $t_p(n, K_{2,k}) = n^p(1 + o(1))$.*

Proof: The lower bound is obtained by considering the complete bipartite graph $K_{a-1,n-a+1}$. It contains no $K_{a,k}$ and obviously has $e_p(K_{a-1,n-a+1}) = (a-1)n^p(1+o(1))$. The upper bound is obtained as follows. If G has n vertices and no $K_{a,k}$ then

$$\sum_{i=1}^{n} \binom{d_i}{k} \le (a-1)\binom{n}{k}$$

since otherwise, by the pigeonhole principle, there would be at least a vertices whose neighborhood intersection contains at least k vertices, and hence there would be a $K_{a,k}$. It follows that $e_k(G) \le (a-1)n^k(1 + o(1))$. In case $a = 2$ we get $e_k(G) \le n^k(1 + o(1))$, so by a trivial convexity argument we get that for $p \ge k$ $e_p(G) \le n^p(1 + o(1))$. $\Box$

6 Concluding remarks and open problems

We start this section with a conjecture regarding $t_p(n, H)$ for graphs H which are non-bipartite:

Conjecture 6.1 *Let H be a graph with chromatic number $r \geq 3$. Then, for all $p \geq 1$, $t_p(n,H) = \left(\frac{r-2}{r-1}\right)^p n^{p+1}(1+o(1))$.*

Note that since Conjecture 6.1 is also stated for $p = 1$, this means that the extremal graphs for all p are determined by the Turán Graph. Theorem 1.1 shows that this conjecture holds when H is a complete graph.

The next conjecture states that the extremal graphs yielding $t_2(n,H)$ are also the extremal graphs yielding $t_p(n,H)$ (assuming, as usual, that n is sufficiently large).

Conjecture 6.2 *Let H be a fixed graph, and let $p > 2$ be an integer. There exists $n_0 = n_0(p,H)$ such that for all $n > n_0$, and for all n-vertex graphs G that are H-free the following holds:*

$$t_p(n,H) = e_p(G) \iff t_2(n,H) = e_2(G).$$

Computing $t_p(n,H)$, or even $t_2(n,H)$ for some specific fixed graphs H seems an interesting open problem. The smallest graph for which we have no exact answer is C_4.

Problem 6.3 *Determine $t_2(n,C_4)$. In particular, is it true that for infinitely many n, $t_2(n,C_4) = e_2(F_n)$ where F_n is the friendship graph.*

Recall that by Proposition 5.1, $t_2(n,C_4) = n^2(1+o(1))$. From Turán Theory we know that the Turán number of C_{2k} for $k > 2$ (cf. [1]) is smaller than that of C_4. This is not the case for $p \geq 2$, since C_{2k} contains P_{2k} and hence $t_p(n,C_{2k}) \geq (k-1)n^p(1+o(1))$. On the other hand, since C_{2k} is a subgraph of $K_{k,k}$ we know by Proposition 5.1 that $t_k(n,C_{2k}) = (k-1)n^k(1+o(1))$. It is interesting to determine what happens for $p \neq k$. We conjecture:

Conjecture 6.4 *For $p > 1$, $t_p(n,C_{2k}) = (k-1)n^p(1+o(1))$.*

References

[1] B. Bollobás, *Extremal Graph Theory*, Academic Press, London, 1978.

[2] O.D. Byer, *Two path extremal graphs and an application to a Ramsey-type problem*, Disc. Math. 196 (1999), 51-64.

[3] Y. Caro and R. Yuster, *Graphs with large variance*, Ars Combinatoria, in press.

[4] D. de Caen, *An upper bound on the sum of squares of degrees in a graph*, Discrete Math. 185 (1998), 245-248.

[5] P. Erdős, *Extremal problems in graph theory*, In: Theory of Graphs and its applications, Proc. Sympos. Smolenice, Prague (1964), 29-36.

[6] P. Erdős, *On the graph theorem of Turán*, Mat. Lapok 21 (1970), 249-251.

[7] R.J. Faudree and R.H. Schelp, *Ramsey type results*. Infinite and finite sets-II (A. Hajnal et al eds.) North-Holland, Amsterdam (1975) 657-665.

[8] Z. Füredi, *On the number of edges of quadrilateral-free graphs*, J. Combin. Theory Ser. B. 68 (1996), 1-6.

[9] D. Olpp, *A conjecture of Goodman and the multiplicities of graphs*, Australian J. Combin. 14 (1996), 267-282.

[10] U.N. Peled, R. Petreschi and A. Sterbini, (n, e)*-graphs with maximum sum of squares of degrees*, J. Graph Theory 31 (1999), 283-295.

[11] J.F. Sacle and M.Wozniak, *The Erdős-Sós Conjecture for graphs without* C_4, J. Combin. Theory, Ser. B 70 (1997), 367-372.

[12] A.F. Sidorenko, *Asymptotic solution for a new class of forbidden* r*-graphs*, Combinatorica 9 (1989), 207-215.