Moments of Graphs in Monotone Families

Zoltán Füredi[1,2] and André Kündgen[3]

[1]DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, ILLINOIS 61801
E-mail: z-furedi@math.uiuc.edu, furedi@renyi.hu

[2]RÉNYI INSTITUTE OF MATHEMATICS OF THE HUNGARIAN ACADEMY
BUDAPEST, P. O. BOX 127, HUNGARY-1364

[3]DEPARTMENT OF MATHEMATICS
CALIFORNIA STATE UNIVERSITY SAN MARCOS
SAN MARCOS, CALIFORNIA 92096
E-mail: akundgen@csusm.edu

Received October 10, 2003; Revised March 25, 2005

Published online 17 August 2005 in Wiley InterScience(www.interscience.wiley.com).
DOI 10.1002/jgt.20119

Abstract: The *k*th *moment* of the degree sequence $d_1 \geq d_2 \geq \ldots d_n$ of a graph G is $\mu_k(G) = \frac{1}{n}\sum d_i^k$. We give asymptotically sharp bounds for $\mu_k(G)$ when G is in a monotone family. We use these results for the case $k = 2$ to improve a result of Pach, Spencer, and Tóth [15]. We answer a question of Erdős [9] by determining the maximum variance $\mu_2(G) - \mu_1^2(G)$ of the degree sequence when G is a triangle-free n-vertex graph.

© 2005 Wiley Periodicals, Inc. J Graph Theory 51: 37–48, 2006

Keywords: *monotone graph family; degree sequence; moments; optimization*

Contract grant sponsor: Hungarian National Science Foundation; Contract grant number: OTKA T 032452, T 037846; Contract grant sponsor: National Science Foundation; Contract grant number: DMS 0140692.

1. INTRODUCTION

A family of graphs is *monotone* if it is closed under taking subgraphs. Many basic families of graphs are monotone: forests, k-colorable graphs, graphs of girth $\geq g$, graphs that can be embedded in a fixed surface, and H-free graphs, that is, graphs that do not contain a fixed subgraph H. In this paper, we study properties of the degree sequence of monotone families of graphs.

P. Erdős [9] proved in 1970 that the degree sequence $(\deg(x_1), \ldots, \deg(x_n))$ of every K_p-free graph G with vertices $\{x_1, \ldots, x_n\}$ can be majorized by a $(p-1)$-partite graph H. This means that $V(H) = V(G)$, $\deg_H(x) \geq \deg_G(x)$ for every vertex, and H is $(p-1)$-colorable. This gives a nice short proof for Turán's theorem. In general, it reduces the problem of determining

$$\max_{|V(G)|=n,\ G \text{ is } K_p\text{-free}} \sum_{1 \leq i \leq n} F(\deg_G(x_i)) \tag{1}$$

for a nondecreasing function F to a much simpler optimization problem on partitions of n into $p-1$ positive integers (see [4, 7]). Erdős also observed that this is not necessarily true for general functions F, or graph parameters that depend on the degree sequence in more complicated ways. As a first point of attack, he proposed the following question:

"Let $G(n)$ be a triangle-free graph with vertices $x_1, \ldots, x_n$. What is the maximum possible variance of the sequence of degrees $v(x_1), \ldots, v(x_n)$, and which graphs achieve that maximum?" We completely solve this problem of Erdős in Section 5, by showing that for $n > 3$, the maximum is achieved only by the unbalanced complete bipartite graph(s) whose number of edges is as close to $n^2/8$ as possible.

Another natural problem is to study the equivalent of (1) for monotone families other than K_p-free graphs. In Section 3, we give a simple bound for a general monotone family in the case when F is nondecreasing. This bound is essentially optimal for such important special cases as the kth moments of the degree sequence. Finally, in Section 4, we apply the results for the second moment we obtained in Section 3 to slightly improve a result on crossing numbers by Pach, Spencer, and Tóth [15].

2. DEFINITIONS

Throughout this paper, $\mathcal{G}$ denotes a monotone family of graphs. For any graph parameter F, we can now define

$$F(n; \mathcal{G}) = \max_{|V(G)|=n,\ G \in \mathcal{G}} F(G).$$

The *Turán graph*, $T_p(n)$, is the complete p-partite graph in which all parts are as equal in size as possible. Thus, with the notation $e(G) = |E(G)|$, Turán's theorem says that

$$e(n; \{K_p\text{-free graphs}\}) = e(T_{p-1}(n)), \text{ achieved only by } T_{p-1}(n).$$

The *degree sequence* of a general graph G with vertex set $\{v_1, \ldots, v_n\}$ is the nonincreasing sequence $d_1 \geq d_2 \geq \cdots \geq d_n$, where $d_i = \deg_G(v_i)$. For every function $F: \mathbf{R}^+ \to \mathbf{R}^+$, we define its average value over the degree sequence of such a graph as

$$F(G) = \frac{1}{n}\sum_{i=1}^{n} F(d_i).$$

For example, for $\mu_k(x) = x^k$, we obtain the kth *moment* (of the degree sequence) of G,

$$\mu_k(G) = \frac{1}{n}\sum_{i=1}^{n} d_i^k.$$

Thus the *average degree* of G is $\mu_1 = \mu_1(G) = \frac{2e(G)}{n}$ and the *variance* of G is

$$\sigma^2(G) = \frac{1}{n}\sum_{i=1}^{n}(d_i - \mu_1)^2 = \mu_2(G) - \mu_1^2.$$

Hence the problem in (1) corresponds to determining $F(n; \{K_p\text{-free graphs}\})$ and the question of Erdős corresponds to determining $\sigma^2(n; \{K_3\text{-free graphs}\})$. In [4, 7], the problem of determining $\mu_k(n; \{K_p\text{-free graphs}\})$ exactly is studied, and $\mu_k(n; \{H\text{-free graphs}\})$ is determined approxiamtely for H nonbipartite. We will be more concerned with the case when $e(n; \mathcal{G}) = o(n^2)$, corresponding to the case when H is bipartite.

An important observation is that the number of edges of a graph in a (nontrivial) monotone family is bounded from above by a nontrivial function of the number of its vertices. If f is any real-valued function, which is defined for the natural numbers, then we say that $\mathcal{G}$ is *f-dense* if every n-vertex graph $G \in \mathcal{G}$ satisfies

$$\frac{e(G)}{n^2/2} \leq f(n).$$

So Turán's theorem, for example, implies that K_p-free graphs are $(1 - \frac{1}{p-1})$-dense.

3. DEGREE VARIATION IN MONOTONE FAMILIES

We start by proving a simple bound on the large degrees of a monotone family:

Lemma 3.1. *Let G be an n-vertex graph in an f-dense monotone family of graphs. If the degree sequence of G is $d_1 \geq d_2 \geq \cdots \geq d_n$, then $d_i \leq 2nf(2i)$, whenever $1 \leq i \leq n/2$.*

Proof. Suppose that $V(G) = \{v_1, v_2, \ldots, v_n\}$ where $d(v_j) = d_j$. Fix $1 \leq i \leq n/2$ and let $A = \{v_1, v_2, \ldots, v_i\}$. Randomly pick a set of i additional vertices $B \subseteq V(G) - A$. Let X count the number of edges induced by $A \cup B$. On the one hand, $\mathrm{E}X \leq \frac{(2i)^2}{2} f(2i)$. On the other hand, $\mathrm{E}X \geq \sum_{j=1}^{i} d_j i/n$, since every edge induced by A is induced by $A \cup B$ with probability $1 \geq i/n + i/n$, and every every edge incident to exactly one vertex in A is induced by $A \cup B$ with probability $i/(n-i) \geq i/n$. Thus

$$2i^2 f(2i) \geq \mathrm{E}X \geq \sum_{j=1}^{i} d_j i/n \geq d_i i^2/n. \qquad \blacksquare$$

Lemma 3.1 enables us to prove the following general bound.

Theorem 3.2. *If F is nondecreasing, and the monotone family $\mathcal{G}$ is f-dense, then*

$$nF(n, \mathcal{G}) \leq 2 \sum_{i=1}^{n/2} F(2nf(2i)).$$

If f is a nonincreasing function on $\mathbf{R}^+$, then furthermore

$$nF(n, \mathcal{G}) \leq 2F(2nf(2)) + \int_2^n F(2nf(x))\, dx.$$

Proof. Let $G \in \mathcal{G}$ be an n-vertex graph achieving $F(n, \mathcal{G})$. Now

$$\begin{aligned} nF(n, \mathcal{G}) &= \sum_{i=1}^{n} F(d_i) \leq 2 \sum_{i=1}^{n/2} F(d_i) \leq 2 \sum_{i=1}^{n/2} F(2nf(2i)) \\ &= 2F(2nf(2)) + 2 \sum_{i=2}^{n/2} F(2nf(2i)). \end{aligned}$$

If $f(x)$ is nonincreasing, then so is $F(2nf(2t))$ and with the substitution $2t = x$, we obtain,

$$2 \sum_{i=2}^{n/2} F(2nf(2i)) \leq 2 \int_1^{n/2} F(2nf(2t))\, dt = \int_2^n F(2nf(x))\, dx. \qquad \blacksquare$$

To see that the bounds in Theorem 3.2 can be quite sharp, we use them to estimate the kth moment of most monotone families.

Theorem 3.3. *Suppose $\mathcal{G}$ is a monotone family of graphs with unbounded maximum degree such that the maximum number of edges of an n-vertex graph in $\mathcal{G}$ is $\Theta(n^{2-\beta})$, for fixed $0 \le \beta \le 1$.*

1. *If $k\beta < 1$, then $\mu_k(n;\mathcal{G}) = \Theta(n^{k-k\beta})$.*
2. *If $k\beta = 1$, then $\Omega(n^{k-1}) \le \mu_k(n;\mathcal{G}) \le O(n^{k-1}\log n)$.*
3. *If $k\beta > 1$, then $\mu_k(n;\mathcal{G}) = \Theta(n^{k-1})$.*

Proof. Since $\mathcal{G}$ must be $cn^{-\beta}$-dense for some $c > 0$ it follows from Theorem 3.2 that

$$n\mu_k(n;\mathcal{G}) \le 2(2nc2^{-\beta})^k + \int_2^n (2ncx^{-\beta})^k\,dx = 2^{1+k-k\beta}(cn)^k + (2nc)^k \int_2^n x^{-k\beta}\,dx.$$

If $k\beta = 1$, then this yields

$$n\mu_k(n;\mathcal{G}) \le 2^k(cn)^k + (2cn)^k(\log n - \log 2) \le Cn^k \log n,$$

as desired. Otherwise, we obtain

$$\begin{aligned} n\mu_k(n;\mathcal{G}) &\le 2^{1+k-k\beta}(cn)^k + \frac{(2cn)^k}{1-k\beta}(n^{1-k\beta} - 2^{1-k\beta}) \\ &= \frac{k\beta}{k\beta - 1} 2^{1+k-k\beta}(cn)^k + \frac{(2c)^k}{1-k\beta} n^{1+k-k\beta}. \end{aligned}$$

For $k\beta < 1$, the first expression is negative and can be dropped to obtain the desired bound. When $k\beta > 1$, the second expression is negative, and we proceed similarly.

For the lower bounds in cases 2 and 3, observe that since $\mathcal{G}$ contains graphs whose maximum degrees go to infinity it must, by monotonicity, contain every star $K_{1,n-1}$. Hence $n - 1 \ge n/2$ implies

$$\mu_k(n;\mathcal{G}) \ge \mu_k(K_{1,n-1}) = \tfrac{1}{n}((n-1)^k + 1^k + 1^k + \cdots + 1^k) \ge n^{k-1}/2^k.$$

Finally, observe that $\mathcal{G}$ contains a graph G on n vertices with $e = e(G) \ge cn^{2-\beta}$. Thus

$$\mu_k(n;\mathcal{G}) \ge \mu_k(G) = \frac{1}{n}\sum_{i=1}^n d_i^k \ge \frac{1}{n}\sum_{i=1}^n \left(\frac{2e}{n}\right)^k \ge (2c)^k n^{k-k\beta}.$$

This establishes the last lower bound. ■

Remark 3.4. One way of reading Theorem 3.3.1 is that the monotonicity of $\mathcal{G}$ ensures that for $k < 1/\beta$, the kth moments are like those of $cn^{1-\beta}$-regular graphs. In other words, the degree sequences of dense graphs in monotone families are in some sense quite regular.

Conjecture 3.5. *In Theorem* 3.3.2*, the lower bound gives the right answer.*

For example, let $\mathcal{G}$ be the family of all graphs that do not contain a 4-cycle. The maximum number of edges of a 4-cycle free graph on n vertices is $\Theta(n^{3/2})$, so we consider the case when $\beta = 1/2$ and $k = 2$. Observe that $\sum d_i^2$ is the number of walks of length 2 in G. Since in a C_4-free graph, there is at most one such walk between any two different vertices, we conclude that $n\mu_2(n;\mathcal{G}) = \sum d_i^2 \leq 2\binom{n}{2} + 2e(G) = \Theta(n^2)$.

A. Sum of Squares in General Graphs

It is crucial in Theorem 3.3 that $\mathcal{G}$ is monotone as is illustrated by the host of results for general graphs collected in this subsection.

There has been considerable research investigating the maximum of the sum of the squares of the degrees $n\mu_2$ for general n-vertex graph on e edges. Frequently, these results are formulated as the maximum number of paths of length 2, $\sum \binom{d_i}{2} = \frac{n}{2}\mu_2 - e$. This problem was investigated, and basically solved, for bipartite graphs by Ahlswede and Katona [2] and by Aharoni [1] in the form of maximum norms of 0-1 matrices.

Recently de Caen [6] gave a short proof that $n\mu_2(G) = \sum d_i^2 \leq e(2e/n - 1 + n - 2)$, but for $e = cn^{2-\beta}$, this yields only that $\mu_2(G) \leq c'n^{2-\beta}$. For further results related to de Caen's bound, see Das [8], Li and Pan [10], Olpp [12], Peled, Petreschi, and Sterbini [13], and for the graphs with maximum μ_2 for n, e fixed, see Byer [5].

Bey [3] gives a generalization of de Caen's bound to hypergraphs. A general bound on μ_k can be found in Székely, Clark, and Entringer [16].

4. CROSSING NUMBERS FOR MONOTONE FAMILIES

Pach, Spencer, and Tóth [15] proved the following result for monotone families, thus answering a question of Simonovits:

Theorem 4.1. *Consider a monotone family $\mathcal{G}$ of graphs, which is $O(n^{\alpha-1})$-dense for some $0 < \alpha \leq 1$. There are constants $c, c' > 0$ such that the crossing number of any graph $G \in \mathcal{G}$ with n vertices and $e \geq cn \log^2 n$ edges is at least*

$$c' \frac{e^{2+1/\alpha}}{n^{1+1/\alpha}}.$$

Moreover, if there are n-vertex graphs in $\mathcal{G}$ with $\Theta(n^{1+\alpha})$ edges, then this bound is asymptotically optimal up to a constant factor.

Pach, Spencer, and Tóth conjectured that this bound holds even for graphs with $e \geq cn$ edges for a suitable constant $c > 0$. Using the results from Section 3, we make some progress on their conjecture.

The *bisection width*, $\mathrm{b}(G)$, of an n-vertex graph G is the minimum number of edges with exactly one endpoint in W, taken over all $W \subset V(G)$ such that $n/3 \le |W| \le 2n/3$. The following result of Pach, Shahroki, and Szegedy [14] is the main tool in the proof of Theorem 4.1:

Theorem 4.2. *If G is a graph on n vertices, then*

$$\mathrm{b}(G) \le 10\sqrt{\mathrm{cr}(G)} + 2\sqrt{n\mu_2(G)}.$$

The proof of the following result is only a slight modification of the proof of Theorem 4.1.

Theorem 4.3. *Consider a monotone family $\mathcal{G}$ of graphs which is $O(n^{\alpha-1})$-dense for some $0 < \alpha < 1/2$. There are constants $c, c' > 0$ such that the crossing number of any graph $G \in \mathcal{G}$ with n vertices and $e \ge cn\log n$ edges is at least*

$$c'\frac{e^{2+1/\alpha}}{n^{1+1/\alpha}}.$$

Proof. Suppose $\mathcal{G}$ is a monotone $An^{\alpha-1}$-dense family for some $A > 0$ and $0 < \alpha < 1/2$. Suppose that $G \in \mathcal{G}$ has n vertices and $e \ge cn\log n$ edges (for a c yet to be determined). Aiming for a contradiction, we assume furthermore that $\mathrm{cr}(G) < c'(e^{2+1/\alpha})/(n^{1+1/\alpha})$, for suitable c'.

We break G into smaller pieces according to the following procedure:

DECOMPOSITION ALGORITHM

Step 0. Let $G^0 = G, G_1^0 = G, M_0 = 1, m_0 = 1$.

Suppose we already have executed Step i, and the resulting graph G^i, consists of M_i components $G_1^i, G_2^i, \ldots, G_{M_i}^i$, each of at most $(2/3)^i n$ vertices. Assume, without loss of generality, that the first m_i components of G^i have at least $(2/3)^{i+1}n$ vertices and the remaining components have fewer. That is, for $1 \le j \le m_i$, we have

$$(2/3)^{i+1}n \le n(G_j^i) \le (2/3)^i n,$$

so that

$$m_i \le (3/2)^{i+1}. \qquad (2)$$

Step $i+1$. If

$$\left(\frac{2}{3}\right)^i < \frac{1}{A^{1/\alpha}} \cdot \frac{e^{1/\alpha}}{n^{1+1/\alpha}},$$

then STOP.

Else for $1 \leq j \leq m_i$ delete $\mathrm{b}(G_j^i)$ edges from G_j^i so that G_j^i falls into two components, each of at most $(2/3)n(G_j^i)$ vertices. Let G^{i+1} denote the resulting graph on n vertices. Clearly, each component of G^{i+1} has at most $(2/3)^{i+1}n$ vertices.

Suppose that the DECOMPOSITION ALGORITHM terminates in Step $k+1$. If $k > 0$, then

$$\left(\frac{2}{3}\right)^k < \frac{1}{A^{1/\alpha}} \cdot \frac{e^{1/\alpha}}{n^{1+1/\alpha}} \leq \left(\frac{2}{3}\right)^{k-1}. \tag{3}$$

We will first show that G^k contains less than $e/2$ edges: The number of vertices of each component of G^k satisfies

$$n(G_j^k) \leq \left(\frac{2}{3}\right)^k n < \frac{1}{A^{1/\alpha}} \cdot \frac{e^{1/\alpha}}{n^{1+1/\alpha}} n = \left(\frac{e}{An}\right)^{1/\alpha}.$$

Since G is from a monotone $An^{\alpha-1}$-dense family, it follows that

$$e(G_j^k) \leq A\frac{n^{1+\alpha}(G_j^k)}{2} < n(G_j^k)\frac{e}{2n}.$$

Thus it follows, as desired, that

$$e(G^k) = \sum_{j=1}^{M_k} e(G_j^k) < \frac{e}{2n}\sum_{j=1}^{M_k} n(G_j^k) = \frac{e}{2}.$$

To obtain a contradiction, it now suffices to show that we deleted at most $e/2$ edges of G to obtain G^k. Using the fact that, for any nonnegative reals $a_1, a_2, \ldots, a_m$

$$\sum_{j=1}^{m} \sqrt{a_j} \leq \sqrt{m\sum_{j=1}^{m} a_j},$$

and (2) we obtain that, for any $0 \leq i < k$,

$$\sum_{j=1}^{m_i} \sqrt{\mathrm{cr}(G_j^i)} \leq \sqrt{m_i \sum_{j=1}^{m_i} \mathrm{cr}(G_j^i)} \leq \sqrt{\left(\frac{3}{2}\right)^{i+1}} \sqrt{\mathrm{cr}(G)} < \sqrt{\left(\frac{3}{2}\right)^{i+1}} \sqrt{c' \frac{e^{2+1/\alpha}}{n^{1+1/\alpha}}}.$$

Using (3) we also obtain

$$\sum_{i=0}^{k-1} \sqrt{\left(\frac{3}{2}\right)^{i+1}} = \sum_{i=1}^{k} \sqrt{\frac{3}{2}}^{i} = \frac{\sqrt{\frac{3}{2}}^{k+1} - \sqrt{\frac{3}{2}}}{\sqrt{\frac{3}{2}} - 1} < 7\sqrt{\left(\frac{3}{2}\right)^{k-1}} \leq 7\sqrt{A^{1/\alpha}\frac{n^{1+1/\alpha}}{e^{1/\alpha}}}.$$

Setting $n_j^i = |V(G_j^i)|$, $k = 2$, and $\beta = 1 - \alpha > 1/2$, it now follows from Theorem 3.3.3 that

$$\sum_{j=1}^{m_i} \sqrt{n_j^i\ \mu_2(G_j^i)} \le \sum_{j=1}^{m_i} \sqrt{n_j^i\ \mu_2(n_j^i, \mathcal{G})} \le \sum_{j=1}^{m_i} c'' n_j^i \le c'' n.$$

Using Theorem 4.2, we see that the total number of edges deleted during the procedure is

$$\begin{aligned}\sum_{i=0}^{k-1}\sum_{j=1}^{m_i} b(G_j^i) &\le 10\sum_{i=0}^{k-1}\sum_{j=1}^{m_i}\sqrt{\mathrm{cr}(G_j^i)} + 2\sum_{i=0}^{k-1}\sum_{j=1}^{m_i}\sqrt{n_j^i\ \mu_2(G_j^i)}\\ &\le 10\sum_{i=0}^{k-1}\sqrt{\left(\frac{3}{2}\right)^{i+1}}\sqrt{c'\frac{e^{2+1/\alpha}}{n^{1+1/\alpha}}} + 2\sum_{i=0}^{k-1} c'' n\\ &< 70\sqrt{A^{1/\alpha}c' e^2} + 2kc'' n\end{aligned}$$

and this expression is at most $e/2$ for $e \ge cn\log n$ provided that c is sufficiently large and c' sufficiently small. ■

A similar argument can be used to show that when $\alpha = 1/2$, then it is sufficient to require $e \ge cn\log^{3/2} n$ for c sufficiently large. We also note that the sharpness of our bounds in Theorem 3.3 suggests that the techniques in [15] alone will not suffice to settle their conjecture.

5. THE VARIANCE OF TRIANGLE-FREE GRAPHS

We now let $\mathcal{G}$ denote the family of triangle-free graphs and let $\sigma^2(n) = \sigma^2(n; \mathcal{G})$ denote the maximum variance $\sigma^2(G)$ of the degree sequence of any triangle-free graph G on n vertices. Hence the question of Erdős mentioned in the introduction is to determine $\sigma^2(n)$ and find all n-vertex $G \in \mathcal{G}$ such that $\sigma^2(G) = \sigma^2(n)$. We do not give an explicit formula for $\sigma^2(n)$, but instead describe the extremal graphs.

We start our investigation by computing $\sigma^2(G)$ when G is complete bipartite:

$$\begin{aligned}\mu_2(K_{k,n-k}) &= \frac{1}{n}\sum_{i=1}^{n} d_i^2 = \frac{1}{n}(k(n-k)^2 + (n-k)k^2) = k(n-k) = e,\\ \mu_1(K_{k,n-k}) &= \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{2e}{n},\\ \sigma^2(K_{k,n-k}) &= \mu_2 - \mu_1^2 = e - \frac{4e^2}{n^2} = \frac{n^2}{16} - \frac{4}{n^2}\left(\frac{n^2}{8} - e\right)^2.\end{aligned} \tag{4}$$

Next, observe that for every positive integer n, one can choose k such that

$$\left|\frac{n^2}{8} - k(n-k)\right| < 0.38n. \tag{5}$$

For example, for $n \geq 10$, one can define k as the closest integer to $n(2-\sqrt{2})/4 \sim 0.1464\ldots n$, and for $3 \leq n \leq 10$ let $k = 1$. Thus (4) and (5) imply

$$\sigma^2(n;\mathcal{G}) \geq \max_{1 \leq k < n} \sigma^2(K_{k,n-k}) > \frac{n^2}{16} - 0.58. \tag{6}$$

Let $\mathcal{B}_n$ denote the set of complete bipartite graphs $K_{k,n-k}$ for which $e = k(n-k)$ is closest to $n^2/8$. $\mathcal{B}_n$ usually consists of a single graph, but an elementary number theoretical argument shows that there are also infinitely many cases when $\mathcal{B}_n = \{K_{k,n-k}, K_{k+1,n-k-1}\}$. Indeed, the equation $n^2/8 - k(n-k) = (k+1)(n-k-1) - n^2/8$ leads to the Pell-type equation

$$(2k+1)^2 - 2\left(\frac{n}{2} - (2k+1)\right)^2 = 1,$$

and all of its integer solutions can be obtained by a simple recurrence, see, e.g., the textbook [11] (page 352). The first few cases are $(n,k,k+1) = (2,0,1)$, $(10,1,2), (58,8,9), (338,49,50), (1970,288,289)\ldots$

Theorem 5.1. *If G is a triangle-free graph on n vertices, then $\sigma^2(G) \leq \sigma^2(n)$. When $n > 3$ equality holds only for the (unbalanced complete bipartite) graphs $G \in \mathcal{B}_n$.*

Proof. Suppose that G is a triangle-free n-vertex graph with $\sigma^2(G) = \sigma^2(n)$. By the preceding observations, it is sufficient to show that G is a complete bipartite graph. Let $d_1 \geq \cdots \geq d_n$ be the degree sequence of G, where $\deg_G(v_i) = d_i$.

First, we give an upper bound for $\mu_2 = \mu_2(G)$: Consider an edge connecting the vertices v_i and v_j. Since G is triangle-free, the neighborhoods $N(v_i)$ and $N(v_j)$ are disjoint, and thus $d_i + d_j \leq n$. Adding this inequality for all edges, we obtain $n\mu_2 = \sum d_i^2 \leq ne$, where $e = |E(G)|$. Thus $\mu_2 \leq e$ and $\mu_1 = 2e/n$ yield

$$\sigma^2(G) = \mu_2 - \mu_1^2 \leq e - \frac{4e^2}{n^2} = \frac{n^2}{16} - \frac{4}{n^2}\left(\frac{n^2}{8} - e\right)^2. \tag{7}$$

Let $K_{k,n-k} \in \mathcal{B}_n$. Since $\sigma^2(K_{k,n-k}) \leq \sigma^2(n) = \sigma^2(G)$, we get from (4), (7), and (5) that

$$\left|\frac{n^2}{8} - e(G)\right| \leq \left|\frac{n^2}{8} - k(n-k)\right| < 0.38n. \tag{8}$$

Define the *deficiency* of the edge v_iv_j by $\varepsilon_{ij} = n - d_i - d_j \geq 0$. Adding the equation $d_i + d_j = n - \varepsilon_{ij}$ for all edges, we obtain $n\mu_2 = ne - \sum \varepsilon_{ij}$ and thus the following improvement of (7)

$$\sigma^2(G) = \mu_2 - \mu_1^2 = e - \frac{1}{n}\left(\sum_{v_iv_j \text{ an edge}} \varepsilon_{ij}\right) - \frac{4e^2}{n^2}$$

$$= \frac{n^2}{16} - \frac{1}{n}\left(\sum \varepsilon_{ij}\right) - \frac{4}{n^2}\left(\frac{n^2}{8} - e\right)^2 \leq \frac{n^2}{16} - \frac{1}{n}\sum \varepsilon_{ij}. \tag{9}$$

Equations (9) and (6) give

$$\sum_{v_iv_j \text{ an edge}} \varepsilon_{ij} < 0.58n. \tag{10}$$

Call an edge v_iv_j *saturated* if $\varepsilon_{ij} = 0$. Since for every nonsaturated edges v_iv_j, we have $\varepsilon_{ij} \geq 1$, it follows from (8) and (10) that there are at least $(n^2/8 - 0.38n) - 0.58n$ saturated edges. This is positive for $n > 7$, so in this case, there is at least one saturated edge, say v_iv_j, with degrees $d_i = \ell$ and $d_j = n - \ell$.

Since G is triangle-free, $N(v_i)$ and $N(v_j)$ are independent sets. Thus G is bipartite with parts $N(v_i)$ and $N(v_j)$. Every vertex in $N(v_j)$ has degree at most ℓ. If all vertices in $N(v_j)$ have degree exactly ℓ, then G is a complete bipartite graph, and we are done. Otherwise, let A denote the set of vertices in $N(v_j)$ of degrees exactly ℓ, and let $A' := N(v_j) \setminus A$. Since $v_i \in A$, we have that A and A' are nonempty.

Similarly, if G is not the complete bipartite graph $K(\ell, n - \ell)$, then we have $B \neq \emptyset$, $B' \neq \emptyset$, where $B := \{y \in N(v_i) : \deg(y) = n - \ell\}$ and $B' := N(v_i) \setminus B$.

Since the vertices of A are joined to all vertices of B' and the vertices of B are joined to all vertices of A', the induced bipartite graphs $G[A, B']$ and $G[B, A']$ are connected and cover all vertices of G. Thus they have at least $n - 2$ edges. None of these edges are saturated, so that (10) yields $n - 2 < 0.58n$. This is a contradiction for $n > 4$.

For $4 \leq n \leq 7$, there are only a few graphs to check. In these cases, stars give the maximum variance, thus finishing the proof of the theorem: When $n = 5, 6, 7$, inequality (5) can be improved to $|n^2/8 - (n-1)| < 0.18n$, which gives an upper bound of $0.14n$ in (10), so that all edges of G are saturated. Similarly, for $n = 4$, at most 1 edge is unsaturated and thus we only need to consider $K_{1,3}$, $K_{2,2}$ and graphs with at most one edge.

The cases $n \leq 3$ are obvious: there are two extrema, the one-edge graph and its complement. Thus we can determine $\sigma^2(n)$ for all n. ■

ACKNOWLEDGMENTS

The authors thank Bruce Richter for suggesting the study of $\mu_2(n; \mathcal{G})$ and its connections to Theorem 4.1. The authors also, thank Radhika Ramamurthi for helpful discussions.

REFERENCES

[1] R. Aharoni, A problem in rearrangements of $(0, 1)$ matrices, Discrete Math 30 (1980), 191–201.

[2] R. Ahlswede and G. O. H. Katona, Graphs with maximal number of adjacent pairs of edges, Acta Mathematica Academiae Scientiarium Hungaricae 32 (1978), 97–120.

[3] C. Bey, An upper bound on the sum of squares of degrees in a hypergraph, Discrete Math 269 (2003), 259–263.

[4] B. Bollobás and V. Nikiforov, Degree powers in graphs with forbidden subgraphs, Electron J Combin 11(1) (2004), R42, 8 p.

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

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

[7] Y. Caro and R. Yuster, Degree powers in graphs with forbidden subgraphs, Electron J Combin 7 (2000), R47, 14 p.

[8] K. C. Das, Maximizing the sum of the squares of the degrees of a graph. Discrete Math 285(1–3) (2004), 57–66.

[9] P. Erdős, On the graph theorem of Turán, Matematikai Lapok 21 (1970), 249–251 (In Hungarian, MR46# 7090).

[10] J. Li and Y. Pan, de Caen's inequality and bounds on the largest Laplacian eigenvalue of a graph, Linear Algebra Appl 328 (2001), 153–160.

[11] I. Niven, H. S. Zuckerman, and H. L. Montgomery, An Introduction to the Theory of Numbers, Fifth edition, John Wiley & Sons, Inc., New York, 1991.

[12] D. Olpp, A conjecture of Goodman and the multiplicities of graphs, Austral J Combin 14 (1996), 267–282.

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

[14] J. Pach, F. Shahroki, and M. Szegedy, Applications of the crossing number, Algorithmica 16 (1996), 111–117.

[15] J. Pach, J. Spencer, and G. Tóth, New bounds on crossing numbers, Discrete Comput Geom 24 (2000), 623–644.

[16] L. A. Székely, L. H. Clark, and R. C. Entringer, An inequality for degree sequences, Discrete Math 103 (1992), 293–300.