

Induced Turán problems: Largest P_m -free graphs with bounded degree

Myung S. Chung*, Tao Jiang†, Douglas B. West‡

September 25, 2005

Abstract

A graph is H -free if it has no induced subgraph isomorphic to H . Let $\text{ex}^*(D; H)$ be the maximum number of edges in an H -free connected graph with maximum degree D ; this is finite if and only if H is a disjoint union of paths. Earlier results include $\text{ex}^*(D; P_4) = D^2$ and the exact computation of $\text{ex}^*(D; 2P_3)$. For fixed $m \geq 6$, we prove that $\text{ex}^*(D; P_m) = \frac{1}{2}D^{m/2} + O(D^{m/2-1})$ when m is even and $\text{ex}^*(D; P_m) = \frac{1}{8}D^{(m+1)/2} + O(D^{(m-1)/2})$ when m is odd (we also find the exact maximum number of vertices in such a graph when m is odd). Finally, we obtain $\text{ex}^*(D; P_5)$ exactly when $D \geq 187$: for example, it equals $\frac{2}{27}D^3 + \frac{7}{18}D^2 + \frac{1}{2}D$ when D is divisible by 3.

1 Introduction

The classical Turán problem in extremal graph theory is to determine $\text{ex}(n; H)$, the maximum number of edges in an n -vertex graph not having H as a subgraph. The same question for forbidden *induced* subgraphs is uninteresting, since a complete graph has no induced subgraph isomorphic to H when H is not complete. To obtain a sensible problem, we fix the maximum degree D and consider only connected graphs.

*9536 Whitecedar Court, Vienna, VA 22181-5420.

†Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, jiangt@muohio.edu. Research supported in part by Miami University CFR faculty summer research grant.

‡Department of Mathematics, University of Illinois, Urbana, IL 61801, west@math.uiuc.edu. Research supported by the National Security Agency under Grant Number MDA904-03-1-0037. The United States Government is authorized to reproduce and distribute reprints.

An H -free graph is a graph having no induced subgraph isomorphic to H . Let $\mathcal{F}(D; H)$ be the family of connected H -free graphs with maximum degree at most D . We study $\text{ex}^*(D; H)$, defined to be $\max_{G \in \mathcal{F}(D; H)} e(G)$, where $e(G)$ is the size (number of edges) of a graph G (we use $n(G)$ for the number of vertices). Since an arbitrarily long path has no cycle and no vertex with degree exceeding 2, $\text{ex}^*(D; H)$ is finite only when H is a linear forest (a disjoint union of paths). We write P_m , C_m , and K_m for the path, cycle, and complete graph with m vertices, respectively. Also, $G + H$ denotes the disjoint union of G and H , and tH denotes the graph consisting of t disjoint copies of H . Thus $\sum_{i=1}^t P_i$ is a linear forest with t components, and tP_m is such a forest whose components all have m vertices.

The value of $\text{ex}^*(D; H)$ is known exactly for some small H . In this paper, we extend this to $H = P_5$ and determine $\text{ex}^*(D; P_m)$ asymptotically for fixed m and large D where $m > 5$. This in turns yields the order of growth of $\text{ex}^*(D; tP_m)$ for fixed m .

We first describe the earlier results for small H . Trivially, $\text{ex}^*(D; P_2) = 0$ and $\text{ex}^*(D; P_3) = \binom{D+1}{2}$. F. Chung, Gyárfás, Tuza, and Trotter [2] showed that $\text{ex}^*(D; 2P_2)$ is $5D^2/4$ when D is even and is $(5D^2 - 2D + 1)/4$ when D is odd. When D is even, the unique extremal graph is $C_5[D/2]$, where $G[t]$ denotes the graph obtained from G by expanding each vertex into an independent set of size t . The graph $C_5[D/2]$ is also $(P_2 + P_3)$ -free. M. Chung [3] proved that it is the largest triangle-free graph in $\mathcal{F}(D; P_2 + P_3)$ and proved that $\text{ex}^*(D; P_2 + P_3) < 2D^2$. She conjectured that $\text{ex}^*(D; P_2 + P_3) = 5D^2/4$ for even D ; this remains open.

Chung and West [4] presented several proofs that $\text{ex}^*(D; P_4) = D^2$; the unique extremal graph is the complete bipartite graph $K_{D, D}$. In [5], they determined $\text{ex}^*(D; P_3 + P_3)$ for $D \geq 8$. The value is $\frac{1}{8}(D^4 + D^3) + O(D^2)$, given exactly by choosing k to maximize $k(D + 1 - k)[\binom{D}{2} + 1] + \binom{k}{2}$ (maximized at $k = \lceil (D + 1)/2 \rceil$). To achieve the bound, start with K_k , add $D + 1 - k$ pendant edges at each vertex, and let each vertex so added belong to a copy of K_D with $D - 1$ other new vertices. The number of edges is as computed above, and the graph is $2P_3$ -free since every induced P_3 has a vertex in the central copy of K_k .

In Section 5 of this paper, we determine $\text{ex}^*(D; P_5)$ for $D \geq 187$. The value is the maximum over k of $k \binom{D+2-k}{2} + \binom{k}{2}$. The maximum occurs at $k = \lceil (D + 1)/3 \rceil$, yielding

$$\text{ex}^*(D; P_5) = \frac{2}{27}D^3 + \frac{7}{18}D^2 + \frac{1}{2}D + \begin{cases} 0, & \text{if } D \equiv 0 \pmod{3}; \\ \frac{1}{27}, & \text{if } D \equiv 1 \pmod{3}; \\ -\frac{1}{9}D + \frac{2}{27}, & \text{if } D \equiv 2 \pmod{3}. \end{cases}$$

To achieve these values, start with kK_{D-k+2} , select one vertex in each component, and add edges so that these vertices induce K_k . This graph is P_5 -free, since an induced path has at most two vertices in a clique (a *clique* is a set of pairwise adjacent vertices).

In general, Chung and West [5] observed that

$$\text{ex}^*(D; H' + P_m) \leq \max\{\text{ex}^*(D; H'), n(H')D^2[\text{ex}^*(D; P_m) + 1]\},$$

as follows. Consider $G \in \mathcal{F}(D; H' + P_m)$. If G is H' -free, then $e(G) \leq \text{ex}^*(D; H')$. Otherwise, the components of the graph obtained by deleting from G the vertices in a copy of H' and their neighbors must be P_m -free, and there are at most $n(H')(D-1)^2$ such components.

When P_m is the largest component of H , the recursive bound implies that $\text{ex}^*(D; H) \in O(D^2 \text{ex}^*(D; P_m))$. For example, $(3q-1)D^2/4 \leq \text{ex}^*(D; qP_2) \leq (2q-2)D^2$ for even D , with the lower bound achieved by $C_{3q-1}[D/2]$. Similarly, generalizing the construction of [5] for $2P_3$ yields $\frac{1}{4}D^2 \text{ex}^*(D; P_m) \leq \text{ex}^*(D; 2P_m) \leq mD^2[\text{ex}^*(D; P_m) + 1]$. Further generalizing it and generalizing the recursive upper bound of [5] to $\text{ex}^*(D; H_1 + H_2)$ yields

$$(t-1)D^2 \text{ex}^*(D; P_m) \leq \text{ex}^*(D; tP_m) \leq (1+o(1))(m\sqrt{t}D^2)^{\lceil \log_2 t \rceil} \text{ex}^*(D; P_m).$$

Our main result determines $\text{ex}^*(D; P_m)$ asymptotically for fixed m . We prove that the construction for $m > 5$ presented originally in [5] is asymptotically optimal. The construction depends on the parity of D ; we describe it next and illustrate it in Figure 1.

Construction 1.1 *Large P_m -free graphs, for $m > 5$.* Let $T_{D,s}$ be a rooted tree in which every non-leaf vertex has $D-1$ children and every leaf has distance s from the root; there are $(D-1)^s$ leaves and $\frac{(D-1)^{s+1}-1}{D-2}$ vertices. Let $T'_{D,s}$ be the graph obtained from $T_{D,s}$ and $(D-1)^s K_D$ by identifying each leaf of $T_{D,s}$ with one vertex of one component of $(D-1)^s K_D$ (using disjoint cliques for distinct leaves). The longest induced path in $T'_{D,s}$ is P_{2s+3} .

When m is even, let $G_{D,m}$ consist of D disjoint copies of $T'_{D,m/2-3}$ plus a central star with D leaves whose leaves are the roots of the copies of $T'_{D,m/2-3}$. The longest induced path in $G_{D,m}$ is P_{m-1} , and the maximum degree is D , so $G_{D,m} \in \mathcal{F}(D; P_m)$. The complete graphs at the leaves contribute $\binom{D}{2}D(D-1)^{m/2-3}$ edges. The large tree contributes $D \frac{(D-1)^{m/2-2}-1}{D-2}$ edges, so $e(G_{D,m}) = \frac{1}{2}D^{m/2} + O(D^{m/2-1})$.

When m is odd, let $G_{D,m}$ consist of $\lfloor (D+1)^2/4 \rfloor$ disjoint copies of $T'_{D,(m-7)/2}$, plus $\lceil (D+1)/2 \rceil$ stars (each having $\lfloor (D+1)/2 \rfloor$ leaves) whose leaves are the roots of the copies of $T'_{D,(m-7)/2}$, plus a copy of $K_{\lceil (D+1)/2 \rceil}$ whose vertices are the centers of the stars. The longest induced paths in G have two central vertices in the central clique and extend $(m-3)/2$ edges farther in each direction. These paths have length $m-2$, so $G_{D,m} \in \mathcal{F}(D; P_m)$. The complete graphs at the leaves contribute $\binom{D}{2} \lfloor (D+1)^2/4 \rfloor (D-1)^{(m-7)/2}$ edges. The rest of the graph contributes $\lfloor (D+1)^2/4 \rfloor \left[\frac{(D-1)^{(m-5)/2} - 1}{D-2} + \binom{\lceil (D+1)/2 \rceil}{2} \right]$ edges, so $e(G_{D,m}) = \frac{1}{8}D^{(m+1)/2} + O(D^{(m-1)/2})$. \blacksquare

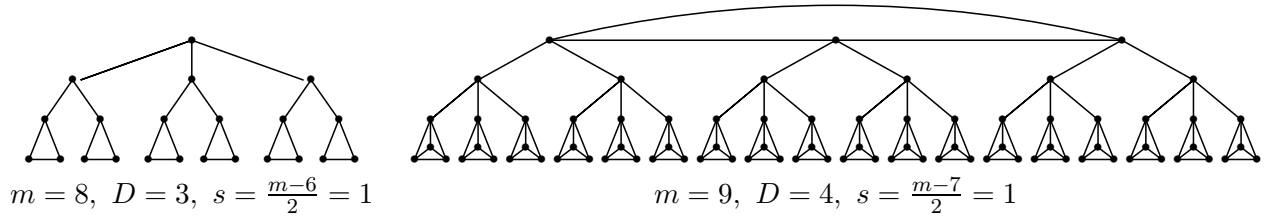


Figure 1: Construction of $G_{D,m}$

In addition to the asymptotic result, in Section 3 we determine the exact maximum number of vertices among graphs in $\mathcal{F}(D; P_m)$ when m is odd, $m > 5$, and $D > m + 20$. The extremal graph is $G_{D,m}$ above. Since every vertex of $G_{D,m}$ has degree at least $D-1$, this result immediately implies the asymptotic result for $\text{ex}^*(D, m)$ when m is odd, but it takes more work than the asymptotic result.

2 Preliminaries for $m > 5$

For $G \in \mathcal{F}(D; P_m)$, summing the vertex degrees yields $e(G) \leq n(G)D/2$. Hence we will bound $e(G)$ by bounding $n(G)$. We write $\Delta(G)$ for the maximum vertex degree of a graph G . Let $N_G(v) = \{u \in V(G) : uv \in E(G)\}$ (the *neighborhood* of v), and $N_G(S) = \bigcup_{v \in S} N_G(v)$. An x, y -*path* is a path with endpoints x and y . We write $d_G(x, y)$ for the distance from x to y in a graph G , and $d_G(x, S) = \min_{v \in S} d_G(x, v)$ when $S \subseteq V(G)$. The maximum distance between vertices in G is the *diameter* of G .

The *depth* of a vertex v in a rooted tree T with root w is $d_T(w, v)$. The depth of T is the maximum depth of a vertex in T . A *breadth-first search (BFS) tree* of G is a rooted spanning

tree T of G such that the depth of each vertex in T equals its distance in G from the root. The *eccentricity* of a vertex v in G is the depth of a BFS tree rooted at v .

Graphs in $\mathcal{F}(D; P_m)$ have diameter at most $m - 2$. Such graphs with maximum degree D may have nearly $(D - 1)^{m-2}$ vertices, roughly the square of the desired bound. To tighten the bound, we start with a BFS tree and then modify it. The notation introduced in this construction will be used throughout the discussion of $\text{ex}^*(D; P_m)$ for $m > 5$.

Construction 2.1 Given $G \in \mathcal{F}(D; P_m)$, choose any $w \in V(G)$. Let p be the eccentricity of w . Let T be a BFS tree with root w , so the length of the w, v -path in T is $d_G(w, v)$. The depth of T is p . Let $w_0 = w$. Let P with vertices w_0, \dots, w_p be a path in T from w to a leaf with maximum depth. For each vertex v , let $T(v)$ be the subtree of T rooted at v . Note that $T(v)$ consists of all vertices whose path to w in T contains v .

Let B be the set of vertices outside P having at least one neighbor on P . Form the tree T' from T as follows: for each $v \in B$, cut the edge from v to its parent in T and make its new parent be the deepest (highest-indexed) vertex of P in $N_G(v)$. These changes are made independently, with no conflict.

For each $v \in B$, let $T'(v)$ be the subtree of T' rooted at v . Let B_l be the set of vertices in B whose highest-indexed neighbor on P (and parent in T') is w_{p+1-l} . Let $V_l = \bigcup_{v \in B_l} V(T'(v))$; note that $V(G) = V(P) \cup V_1 \cup \dots \cup V_p$. See Figure 2. ■

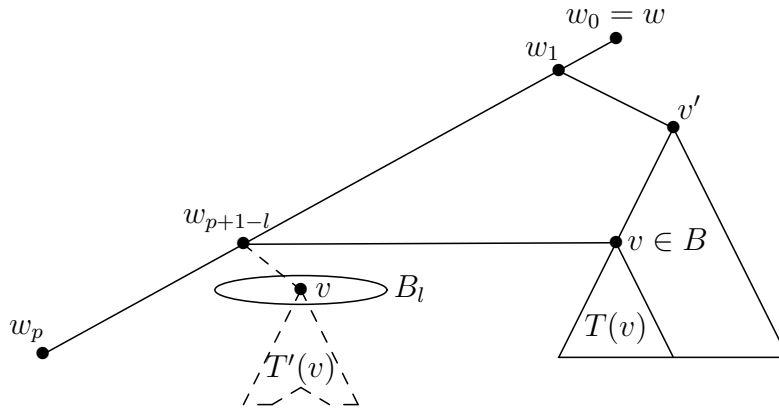


Figure 2: Construction of T'

When x and y are vertices on a path P , we write $P[x, y]$ for the x, y -path along P .

Lemma 2.2 *In the setting of Construction 2.1, let v be a vertex of B_l . If s denotes the depth of $T'(v)$, then $s \leq \min\{l, (m-2) - l\}$. In particular, $s \leq \lfloor m/2 \rfloor - 1$. Furthermore, if $s = \lfloor m/2 \rfloor - 1$, then $l = \lfloor m/2 \rfloor - 1$ or $l = \lceil m/2 \rceil - 1$.*

Proof. Since $v \in B_l$, the parent of v in T' is w_{p+1-l} . Note that l is the length of the path $P[w_p, w_{p+1-l}] \cup w_{p+1-l}v$. Let P' denote this path; it is an induced path by definition of B_l .

Let r be the depth of v in the original tree T , so $d_G(w, v) = r$. Since T has depth p , the depth of $T(v)$ is at most $p - r$. The tree $T'(v)$ is a subtree of $T(v)$; any subtrees of $T(v)$ rooted at other vertices of B were removed. Hence $s \leq p - r$. Since $d_G(w, v) = r$ and $d_G(w, w_p) = p$, the triangle inequality yields $d_G(v, w_p) \geq p - r$. Since P' is a v, w_p -path of length l , we have $d_G(v, w_p) \leq l$, and thus $s \leq l$.

Since $T'(v)$ has depth s , in $T'(v)$ there is a path Q of length s starting at v . Since paths from v in $T'(v)$ are shortest paths joining their endpoints, Q is an induced path. Furthermore, since v is the only vertex of Q in B , the path $Q \cup P'$ is an induced path of length $s + l$. Since G has no induced P_m , we have $s + l \leq m - 2$.

We have shown that $s \leq \min\{l, (m-2) - l\}$. This yields $s \leq \lfloor (m-2)/2 \rfloor = \lfloor m/2 \rfloor - 1$. Furthermore, if $s = \lfloor m/2 \rfloor - 1$, then $l = \lfloor m/2 \rfloor - 1$ or $l = \lceil m/2 \rceil - 1$. ■

Recall that diameter $m - 2$ yields a trivial upper bound of $O(D^{m-2})$ on $\text{ex}^*(D; P_m)$. Lemma 2.2 enables us to reduce the upper bound to within one factor of D of the lower bound from $G_{D,m}$ (Construction 1.1).

Lemma 2.3 $\text{ex}^*(D; P_m) \leq D^{\lceil m/2 \rceil + 1} + O(D^{\lceil m/2 \rceil})$.

Proof. Consider $G \in \mathcal{F}(D; P_m)$, with notation as in Construction 2.1. For $v \in B_l$ or $v \in B_{(m-2)-l}$, where $l \leq m/2 - 1$, Lemma 2.2 implies that $T'(v)$ has depth at most l . When $T'(v)$ has depth s , the computation of Construction 1.1 yields $n(T'(v)) \leq \frac{(D-1)^{s+1} - 1}{D-2}$. Since $B_l \subseteq N_G(w_{p+1-l})$, we have $|B_l| < D$ (in fact, $|B_l| \leq D - 2$ when $2 \leq l \leq p$). We conclude that $|V_l| < D^{s+1} + O(D^s)$, where $s = \min\{l, (m-2) - l\}$. Summing over l yields $n(G) \leq 2D^{\lceil m/2 \rceil} + O(D^{\lceil m/2 \rceil - 1})$. Multiplying $n(G)$ by $D/2$ gives the upper bound on $e(G)$. ■

We will eliminate the extra factor of D and determine the coefficient on the leading term. The proof of Lemma 2.3 allows us to ignore all vertices in $T'(v)$ when $T'(v)$ has depth at

most $\lceil m/2 \rceil - 3$; what remains is A , where $A = \bigcup_{l=\lceil m/2-2 \rceil}^{\lceil m/2 \rceil} V_l$. From Lemma 2.2, we have $n(G - A) = O(D^{\lceil m/2 \rceil - 2})$. Therefore, it suffices to improve the bounds on $|A|$. The next two sections are devoted to this, in the two cases of odd and even m .

We begin with a technical definition.

Definition 2.4 In a graph H , a vertex set S is k -distant from a vertex set C if $d(y, S) = k$ for all $y \in C$. If S is k -distant from C , and there is a vertex $y' \in C$ such that $d(y', y) = k$ and $d(y', v) > k$ for all v in some subset S_0 of S , then we say that y' *requires* y relative to S_0 . When $S_0 = S - \{y\}$, we say simply that y' *requires* y .

When we obtain a vertex that requires y , usually we denote it by y^* . The strongest restriction of the sort in Definition 2.4 is when y^* requires y . The weakest restriction is relative to a single vertex, as in “ y' requires y relative to $\{x\}$ ”. We will need the general form of the restriction in Lemma 4.6. Meanwhile, if S is a minimal set that is k -distant from a set C and has distinct vertices x and y , then the minimality of S guarantees the existence of x^* requiring x and y' requiring y relative to $\{x\}$.

When Q and Q' are disjoint subgraphs of a graph H , let $E(Q, Q')$ denote $\{uv \in E(H) : u \in V(Q), v \in V(Q')\}$. Since we will apply the next lemma many times, we illustrate its scenario in Figure 3 for easy reference.

Lemma 2.5 *Let S be a set that is k -distant from a set C in a graph H , and let x and y be distinct vertices of S . If there exist $x^*, y' \in C$ such that x^* requires x and y' requires y relative to $\{x\}$, and Q and Q' are shortest x, x^* - and (y, y') -paths, respectively, then*

- (1) Q and Q' are disjoint.
- (2) if $uv \in E(Q, Q')$, then $Q[x, u]$ and $Q'[y, v]$ have equal length.
- (3) if $xy \notin E(H)$ but $E(Q, Q') \neq \emptyset$, then $d(y', x_1) = k$, where x_1 neighbors x on Q .

Proof. (1) If Q and Q' share a vertex, then there is an x, y' -path or a y, x^* -path of length at most k . The former contradicts $d(y', x) > k$, and the latter contradicts $d(x^*, y) > k$.

(2) Let Q visit x_0, \dots, x_k in order, with $x_0 = x$ and $x_k = x^*$. Similarly let $\{y_0, \dots, y_k\} = V(Q')$, with $y_0 = y$ and $y_k = y'$. If $x_i y_j \in E(H)$ with $i < j$, then $Q[x_0, x_i] \cup x_i y_j \cup Q'[y_j, y_k]$ is an x, y' -path of length at most k . If $i > j$, then $Q'[y_0, y_j] \cup y_j x_i \cup Q[x_i, x_k]$ is a y, x^* -path of length at most k . As before, both cases are forbidden.

(3) If $xy \notin E(H)$ but $E(Q, Q') \neq \emptyset$, then $x_q y_q \in E(H)$ for some $q \geq 1$. Now $Q[x_1, x_q] \cup x_q y_q \cup Q'[y_q, y_k]$ is an x_1, y' -path of length k . Since $d(y', x) > k$ and $xx_1 \in E(H)$, it follows that $d(y', x_1) = k$. ■

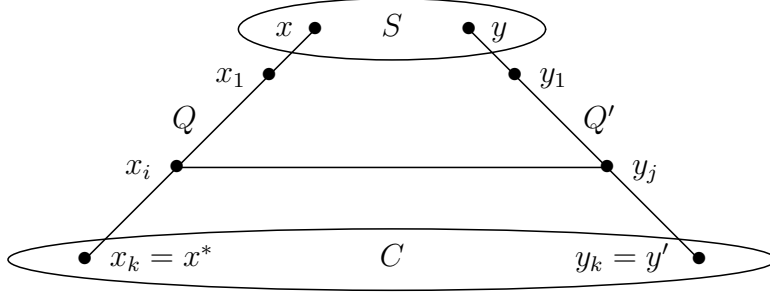


Figure 3: Paths Q and Q'

3 The Upper Bound for Odd m

We continue to use the notation of Construction 2.1 to discuss graphs in $\mathcal{F}(D; P_m)$. When m is odd, the remark after Lemma 2.3 allows us to restrict our attention to the descendants of two consecutive vertices on the path P . For the asymptotic result with m odd, it suffices to bound the sizes of the deepest levels of those parts of T' .

Theorem 3.1 *If $m > 5$ and m is odd, then $\text{ex}^*(D; P_m) = \frac{1}{8}D^{(m+1)/2} + O(D^{(m-1)/2})$.*

Proof. Construction 1.1 gives the lower bound. Now consider $G \in \mathcal{F}(D; P_m)$, with notation as in Construction 2.1. As remarked after Lemma 2.3, it suffices to show that $|A| \leq \frac{1}{4}D^{\lfloor m/2 \rfloor} + O(D^{\lfloor m/2 \rfloor - 1})$, where $A = V_{\lfloor m/2 \rfloor - 1} \cup V_{\lfloor m/2 \rfloor}$. Let H be the subgraph of G induced by A .

Let $k = \lfloor m/2 \rfloor - 1$ and $B' = B_k \cup B_{k+1}$. For $v \in B'$, Lemma 2.2 implies that $T'(v)$ has depth at most k . Thus all vertices of A are within distance k of B' in H . Let $C = \{v \in A: d_H(v, B') = k\}$. Since $B' \subset N_G(w_{p+1-k}) \cup N_G(w_{p-k})$, we have $|B'| < 2D$. Also, $A - C$ is covered by trees rooted in B' with depth at most $k - 1$. Hence $|A - C| < 2D^k + O(D^{k-1})$. Since again $e(G) \leq n(G)D/2$, it thus suffices to prove that $|C| \leq \frac{1}{4}D^{k+1} + O(D^k)$.

Let S be a minimal subset of B' that is k -distant from C . Choose $x \in S$ having the fewest neighbors in S . Let $S_1 = N_H(x) \cap S$ and $S_2 = S - S_1$, so $x \in S_2$. Let $C_1 = \{v \in$

$C: d_H(v, S_1) = k\}$, and let $C_2 = \{v \in C: d_H(v, S_2) = k\}$. Note that $C = C_1 \cup C_2$, though C_1 and C_2 are not necessarily disjoint. To prove the desired bound on $|C|$ and complete the proof, it suffices to prove that $|C_1| \leq \frac{1}{4}D^{k+1}$ and that $|C_2| \leq 2D^k + O(D^{k-1})$.

Claim 1. $|C_1| \leq \frac{1}{4}D^{k+1}$. Let $t = |S_1|$. By our choice of x , each vertex in S_1 has at most $D-t$ neighbors outside S (one of which is in P). Since S is k -distant from C , paths of length k from S_1 to C_1 have no internal vertices in S . Hence fewer than $t(D-t)$ vertices can be reached after one step on a path of length k from S_1 to C_1 , and $t(D-t) \leq D^2/4$. Since each path extends in at most $D-1$ ways at each step until reaching C_1 , we have $|C_1| \leq \frac{1}{4}D^2D^{k-1}$.

Claim 2. $|C_2| \leq 2D^k + O(D^{k-1})$. By the minimality of S , some vertex x^* of C requires x . Let Q be an x, x^* -path of length k in H . For $y' \in C_2 - \{x^*\}$, there exists $y \in S_2$ such that H has a y, y' -path Q' of length k .

If $d_H(y', x) > k$, then $y \neq x$ and y' requires y relative to x . By Lemma 2.5, Q and Q' are disjoint. Also $xy \notin E(H)$, since x has no neighbor in S_2 . Let P' be a shortest x, y -path via P (of length 2 or 3). If $E(Q, Q') = \emptyset$, then $Q \cup P' \cup Q'$ is an induced path of length $m-1$ or m (since $m-1 = 2k+2$). Hence $E(Q, Q') \neq \emptyset$, and now Lemma 2.5 yields $d(x_1, y') = k$.

Thus if $y' \in C_2$, then $d_H(x, y') = k$ or $d_H(x_1, y') = k$, so C_2 is covered by the leaf sets of two trees with depth k . Such a tree has fewer than D^k leaves, so $|C_2| \leq 2D^k + O(D^{k-1})$. ■

We have studied $\text{ex}^*(D; P_m)$ by bounding the number of vertices and then multiplying by $D/2$. If the extremal graphs are D -regular, then the problems are equivalent. However, the example $G_{D,m}$ is not D -regular; most vertices of $G_{D,m}$ have degree $D-1$. Although Theorem 3.1 solves the problem asymptotically for odd m , the irregularity leaves the possibility that another graph in $\mathcal{F}(D; P_m)$ with fewer vertices has more edges.

We show next that $G_{D,m}$ has the most vertices among graphs in $\mathcal{F}(D, m)$ when m is odd and D is large. Since $G_{D,m}$ has minimum degree $D-1$, this also implies Theorem 3.1. We proved Theorem 3.1 separately because the proof motivates our approach for even m and because Theorem 3.2 takes more work. We maintain our earlier notation.

Theorem 3.2 *If m is odd and at least 7, and $D \geq m+20$, then the graph $G_{D,m}$ of Construction 1.1 has the most vertices among graphs in $\mathcal{F}(D; P_m)$.*

Proof. Recall that B_l consists of the vertices in B whose highest-indexed neighbor on P is w_{p+1-l} . Fix $k = (m - 1)/2 - 1$, and let $a = w_{p+1-k}$ and $b = w_{p-k}$. Let $B' = N_G(a) \cup N_G(b)$.

For $v \in B_l$, by Lemma 2.2 $T'(v)$ has depth at most l when $l < m/2 - 1$ and at most $m - 2 - l$ when $l > m/2 - 1$. Hence for $l < m/2 - 1$, all of V_l is within distance $k + 1$ of a (since $l + 1 + k - l = k + 1$), and for $l > m/2 - 1$, all of V_l is within distance k of b (since $m - 2 - l + 1 + l - (k + 1) = k$). Hence each vertex outside B' is within distance k of B' .

Call a subset of B' *good* if every vertex outside B' can reach it within distance k . Among all good subsets of B' , choose S to be one such that

- (1) $|S|$ is smallest,
- (2) subject to (1), $|S \cap \{a, b\}|$ is largest, and
- (3) subject to (1) and (2), S induces the fewest components in G .

Case 1: S is a clique. For $m \geq 9$, each vertex in G has distance at most k from S , so the number of vertices is maximized by growing the largest possible tree of depth k from each vertex of S , subject to each vertex of S having $|S| - 1$ neighbors in S . The construction of $G_{D,m}$ does this, since each root vertex in S gets $D - |S| + 1$ children and each vertex outside S before the k th level gets $D - 1$ children.

Let $r = |S|$. The construction of $G_{D,m}$ is completed by choosing r to be $\lceil (D + 1)/2 \rceil$. To maximize the number of vertices, we choose r to maximize $r + r(D + 1 - r) \sum_{i=1}^k (D - 1)^{i-1}$. This is equivalent to maximizing $r(D + 1 + \epsilon - r)$, where $\epsilon = 1 / \sum_{i=1}^k (D - 1)^{i-1}$. Since $\epsilon < 1$, the optimal choice again is $r = \lceil (D + 1)/2 \rceil$.

For $m = 7$ (and hence $k = 2$), every vertex is within distance k of S except possibly if $S \subseteq N(a)$ or $S \subseteq N(b)$. In this case, $S \cup \{a\}$ or $S \cup \{b\}$ is a clique S' such that every vertex is within distance k of S' , and the above argument suffices again.

Case 2: S is not a clique. Here we show numerically that G has fewer vertices than $G_{D,m}$. Let C be the set of vertices outside B' having distance k from S . Since $|S| \leq 2D$, there are fewer than $2D(\sum_{i=0}^{k-1} (D - 1)^i)$ vertices within distance $k - 1$ of S . Since $(D - 1)^r + D^{r-1} < D^r$ when $r \geq 2$ and $D \geq 2$, the inside sum is less than D^{k-1} (by induction on k). Thus $|V(G) - (B' \cup C)| < 2D^k$. Also $|B'| \leq 2D$. If $|C| < 3D^k$, then $|V(G)| < 5D^k + 2D$. However, $|V(G_{D,m})| = \left\lfloor \frac{(D+1)^2}{4} \right\rfloor \frac{(D-1)^{k-1}}{D-2} + \lceil \frac{D+1}{2} \rceil$, so $|V(G_{D,m})| > (D - 1)^{k-1} D^2 / 4$. Thus $(D - 1)^{k-1} \geq 20D^{k-2}$ suffices, and this inequality holds when $D \geq m + 20$.

To prove that $|C| < 3D^k$, it suffices by the computation above to find a set of three vertices such that every vertex of C is within distance k of this set. Since S was chosen to be a minimal good subset of B' rather than a minimal subset that is k -distant from C , it takes a bit of work to show that for all $x \in S$ there is a vertex $x^* \in C$ that requires x in the sense of Definition 2.4. This is our first task; we use (1), (2), and (3) from the choice of S .

By (1), some vertex outside B' has distance more than k from $S - \{x\}$. Let x^* be such a vertex that is farthest from x . Since S is good, $d(x^*, x) \leq k$. Equality will put $x^* \in C$. Suppose to the contrary that $d(x^*, x) < k$. Since x^* has distance at most k from every neighbor of x , $N(x) \cap S = \emptyset$. Next, if $x \notin \{a, b\}$ but x has a neighbor in $\{a, b\}$, then substituting this neighbor for x in S yields a good subset of B' with the same size as S but more elements in $\{a, b\}$, contradicting (2). Thus $x \in \{a, b\}$. Now $S - \{x\}$ is contained in the neighborhood of the element other than x in $\{a, b\}$. Replacing x with that element yields a good subset of B' with the same size as S and the same size intersection with $\{a, b\}$, but fewer components in its induced subgraph, contradicting (3). Thus $d(x^*, x) = k$, and x^* requires x .

Now, since S is not a clique, we may fix x to be a vertex of S having a nonneighbor in S . Let x^* be a vertex that requires x , and let Q be a shortest x, x^* path. Let y be a nonneighbor of x in S , let y^* be a vertex that requires y , and let Q' be a shortest y, y^* -path. Note that y^* requires y relative to x , meaning that $d(y^*, y) = k$ and $d(y^*, x) > k$. Figure 4 shows the current situation and its further development.

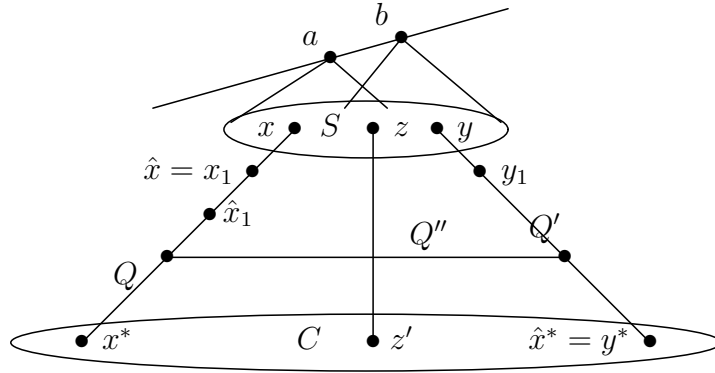


Figure 4: Excluding a non-clique “center”

The set S is k -distant from C , so we can invoke Lemma 2.5. First, Q and Q' are disjoint. If $E(Q, Q') = \emptyset$, then $Q \cup Q'$ together with a shortest x, y -path through $\{a, b\}$ is an induced

path, since the choice of x^* and y^* implies that no vertex of $Q \cup Q'$ other than x or y has a neighbor in B' . The length of such a path would be at least $2k + 2$, which equals $m - 1$. The contradiction yields $E(Q, Q') \neq \emptyset$. Now Lemma 2.5 yields $d(y^*, \hat{x}) = k$, where \hat{x} is the neighbor of x on Q . The choice of x^* implies that \hat{x} has no neighbor in S other than x .

For every $z' \in C$, there is a vertex $z \in S$ with $d(z', z) = k$. If $d(z', x) > k$ and $z \notin N(x)$, then the argument above yields $d(z', \hat{x}) = k$. For the case $z \in N(x)$, we need a third vertex.

Rename the vertex y^* introduced earlier as \hat{x}^* , noting that $d(y^*, \hat{x}) = k$. Since y^* requires y , we have $d(\hat{x}^*, S - \{y\}) > k$. If $d(z', \{x, \hat{x}\}) > k$, then z' requires z relative to \hat{x} . Let \hat{Q} be a shortest \hat{x}, \hat{x}^* -path, and let \hat{x}_1 be the neighbor of \hat{x} on \hat{Q} .

Let Q' be a shortest z, z' -path. To apply Lemma 2.5, we now use the set $\{\hat{x}, z\}$ that is k -distant from the set $\{\hat{x}^*, z'\}$. By Lemma 2.5, Q and Q' are disjoint. If $E(Q, Q') = \emptyset$, then $Q \cup Q'$ together with a shortest \hat{x}, z -path through $\{x, a, b\}$ is an induced path, since $z \notin N(\hat{x})$, and x has no neighbor in Q' except possibly z , and no vertex of $Q \cup Q'$ except z has a neighbor in $\{a, b\}$. The length of this path is at least $2k + 2$, again a contradiction. Hence $E(Q, Q') \neq \emptyset$. Now Lemma 2.5(3) yields $d(z', \hat{x}_1) = k$.

We have proved that every vertex of C is within distance k of $\{x, \hat{x}, \hat{x}_1\}$, as desired. ■

4 The Upper Bound for Even m

The even case is more difficult, because we must consider $V_{m/2-2}$, $V_{m/2-1}$, and $V_{m/2}$. Eliminating the extra factor of D actually requires only the first lemma of this section, but obtaining the correct coefficient on the leading term is much harder.

Throughout this section, we consider $G \in \mathcal{F}(D; P_m)$ for m even and $m > 5$, with notation as in Construction 2.1 for various parts of G . In addition, as in the proof of Theorem 3.1, let $B' = B_{m/2-2} \cup B_{m/2-1} \cup B_{m/2}$, and let $A = V_{m/2-2} \cup V_{m/2-1} \cup V_{m/2}$. By the remark after Lemma 2.3, it suffices to show that $|A| \leq D^{m/2-1} + O(D^{m/2-2})$.

Fix $k = m/2 - 2$, and let $a = w_{p+1-k}$, $z = w_{p-k}$, and $b = w_{p-1-k}$. Thus a and b are the neighbors of z on P , and these three vertices are the parents in T' of the vertices in B' . Let F be the subgraph of G induced by $\{a, z, b\} \cup A$, and let $H = F - (V_{m/2-2} \cup \{a\})$.

The first lemma brings the upper bound to within a factor of 3 of the desired value. We

find a set W of size at most three such that all of A is within distance $m/2 - 1$ of it.

Lemma 4.1 *There is a set W , equal to $\{a, z\}$ or to $\{a, z, u\}$ for some $u \in N_H(z)$, such that all of A is within distance $m/2 - 1$ of W .*

Proof. If $v \in B_{m/2-2}$, then $T'(v)$ has depth at most $m/2 - 2$, by Lemma 2.2. Hence each vertex in $V_{m/2-2}$ has distance at most $m/2 - 1$ from a .

If $d_H(v, z) \leq m/2 - 1$ for all $v \in V(H)$, then the claim holds with $W = \{a, z\}$. If $d_H(v, z) = m/2 + 1$ for some $v \in V(H)$, then $P[w_p, z] \cup Q$ is an induced path of length $m - 1$, where Q is a shortest z, v -path in H (since in $V(H)$, only z has a neighbor in $P[a, w_p]$).

We may therefore assume that z has eccentricity $m/2$ in H (since H is connected). Let $C = \{v: d_H(v, z) = m/2\}$. If $v \in C$, then $d_H(v, N_H(z)) = m/2 - 1$. Thus $N_H(z)$ is l -distant from C , where $l = m/2 - 1$. Let S be a minimal subset of $N_H(z)$ that is l -distant from C . If $|S| \geq 2$, then choose distinct $x, y \in S$. (See Figure 5, where x or y may equal b .)

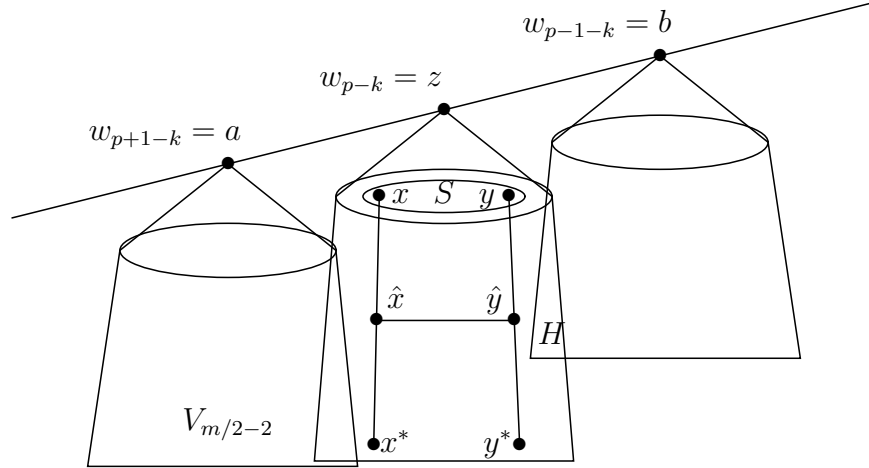


Figure 5: Finding W

Let x^* and y^* be vertices in C that require x and y , respectively. Let Q be a shortest x, x^* -path and Q' a shortest y, y^* -path in H . By Lemma 2.5, Q and Q' are disjoint. If $E(Q, Q') = \emptyset$, then $Q \cup xz \cup zy \cup Q'$ is an induced path of length m . Hence there exists $\hat{x}\hat{y} \in E(Q, Q')$, with $Q[x, \hat{x}]$ and $Q'[y, \hat{y}]$ of equal length. If the only such edge is xy , then $Q \cup xy \cup Q'$ is an induced path of length $m - 1$. Otherwise, for such an edge with $\hat{x}\hat{y} \neq xy$, the path $P[w_p, z] \cup zx \cup Q[x, \hat{x}] \cup \hat{x}\hat{y} \cup Q'[\hat{y}, y^*]$ is an induced path of length $m - 1$.

The contradiction yields $|S| = 1$, and we set $W = \{a, z, u\}$, where $S = \{u\}$. ■

Since at most $D^{m/2-1} + O(D^{m/2-2})$ vertices can be reached within distance $m/2 - 1$ of any vertex, the existence of W yields $|A| \leq 3D^{m/2-1} + O(D^{m/2-2})$. To reduce the leading term by a factor of 3, we must bound the number of vertices in A at maximum distance from W . Let $C = \{v \in A: d_F(v, W) = m/2 - 1\}$. Since $|W| \leq 3$ and $\Delta(F) \leq D$, we have $|A - C| \leq 3D^{m/2-2} + O(D^{m/2-3})$. It thus suffices to show that $|C| \leq D^{m/2-1} + O(D^{m/2-2})$. Presently $|C| < 3D^{m/2-1}$, and the rest of this section is devoted to saving the factor of 3.

Note that $N_F(W) - W$ is k -distant from C , where $k = m/2 - 2$. Henceforth let S be a fixed minimal subset of $N_F(W) - W$ that is k -distant from C . Let $S_a = N(a) \cap S$ and $S_z = N(z) \cap S$. If $W = \{a, z, u\}$, then let $S_u = N(u) \cap S$; otherwise, let $S_u = \emptyset$.

Say that a set U is *semi- k -distant* from C if every vertex of C has distance k or $k - 1$ from U . A set obtained from S by adding some of its neighbors is semi- k -distant from C .

Lemma 4.2 *If a set U is semi- k -distant from C in F , and $|N_F(U) - S| \leq D^2 + O(D)$, then $|C| \leq D^{m/2-1} + O(D^{m/2-2})$.*

Proof. Recall that $k = m/2 - 2$. Since S is k -distant from C , a path of length at most k from U to C cannot start along an edge to a vertex of S . It must move to a vertex in $N_F(U)$. From each such vertex, at most $(D - 1)^{k-1}$ vertices of C can be reached in $k - 1$ additional steps (plus a lower-order term for vertices one step closer). Hence $|C| \leq D^{k+1} + O(D^k)$. ■

In most cases, we will find such a set U to obtain the desired bound on $|C|$. For example, if one of $\{S_a, S_z, S_u\}$ contains the others, then $|S| \leq D$ and $|N(S)| < D^2$, and Lemma 4.2 applies with $U = S$. Therefore, we may assume none of S_a, S_z, S_u contains the other two.

Many of the cases we consider below are eliminated by constructing induced paths of length at least $m - 1$. The stereotypic argument is summarized in the next lemma. This continues the discussion of Lemma 2.5 in the even case, where $k = m/2 - 2$.

Lemma 4.3 *Let x and y be distinct vertices of S . Let x^* be a vertex of C that requires x , let S_0 be a subset of $S - \{y\}$ containing x , and let y' be a vertex of C that requires y relative to S_0 . If $E(Q, Q') = \emptyset$, where Q and Q' are shortest x, x^* - and y, y' -paths, respectively, then there is no induced x, y -path of length at least 3 through $W \cup S_0$.*

Proof. Both Q and Q' have length k . By the choice of x^* and y' and W , the vertices of Q and Q' have no neighbors in $W \cup S_0$, except for the edge incident to x on Q and edges from x and y to $W \cup S_0$. Therefore, if P' is a path as described, then $Q \cup P' \cup Q'$ is an induced path. Its length is at least $2k + 3$, which equals $m - 1$ and is forbidden. ■

In order to apply Lemma 4.2, we must limit the outside neighbors of some subset of S . The next lemmas further this aim by forcing edges within S . Whenever we introduce distinct vertices x and y in S and a set $S_0 \subseteq S - \{y\}$ that contains x , automatically we have x^*, y', Q, Q' defined as in Lemma 4.3. If we do not specify S_0 , then any such S_0 may be used. Often $S_0 = \{x\}$ suffices, but sometimes we will specify a larger set whose members must have distance more than k from y' . Always such a vertex y' exists, since some vertex requires y , and such a vertex automatically also requires y relative to S_0 . We also name the vertices of Q and Q' as x_0, \dots, x_k and y_0, \dots, y_k in order, with $x_0 = x$ and $y_0 = y$.

Here and occasionally later we use that a and u are not adjacent (if u exists). This holds because all neighbors of a in F were discarded to form H , and u is chosen from $N_H(z)$.

Lemma 4.4 *Let x be a vertex of $S_u - (S_a \cup S_z)$.*

- 1) *If $y \in S - \{x\}$ and $E(Q, Q') \neq \emptyset$, then $xy \in E(G)$.*
- 2) *If $y \in (S_a \cup S_z) - S_u$, then $xy \in E(G)$.*

Proof. (1) By Lemma 2.5(2), $x_i y_i \in E(G)$ for some i . If $i > 0$, then let $P' = ux \cup Q[x, x_i] \cup x_i y_i \cup Q'[y_i, y']$; this is an induced path of length $k + 2$. By the construction of T' , within A only u and vertices of S can have neighbors in $P[z, w_p]$, and only a and z can be those neighbors. On P' , only $x \in S$. Since $x \notin S_a \cup S_z$, x has no neighbor in $P[z, w_p]$. Also $u \notin N(a)$. Now $P' \cup uz \cup P[z, w_p]$ is an induced path of length $2k + 3$, a contradiction. Thus $i = 0$, and $xy \in E(G)$.

(2) By the choice of u and the hypotheses here, $xz, xa, ua, uy \notin E(G)$. Hence $\{y, z, u, x\}$ or $\{y, a, z, u, x\}$ (depending on whether $y \in S_z$) induces a path P' . This violates Lemma 4.3 if $E(Q, Q') = \emptyset$, so $E(Q, Q) \neq \emptyset$, and (1) yields $xy \in E(G)$. ■

Lemma 4.5 *If $y \in (S_a \cup S_z) \cap S_u$, then $N(y)$ contains $S_u - (S_a \cup S_z)$ or $(S_a \cup S_z) - S_u$.*

Proof. Let $S_0 = (S_a \cup S_z) - S_u$. If the claim fails, then there exist $x \in S_u - (S_a \cup S_z)$ and $v \in S_0$ such that $yx, yv \notin E(G)$. By Lemma 4.4(2), $xv \in E(G)$.

Since $xy \notin E(G)$, Lemma 4.4(1) implies $E(Q, Q') = \emptyset$. Since v and y are nonadjacent, and each has at least one neighbor in $\{a, z\}$, we have an induced v, y -path of length 2 or 3 through $\{a, z\}$. Since we have $a, z, v \notin N(x)$, adding the edge vx completes an induced x, y -path P' through $W \cup \{v\}$ of length at least 3, which is forbidden by Lemma 4.3. ■

We seek a set that is semi- k -distant from C and is smaller than S . We want to replace a portion of S with a single vertex to obtain U for which we can prove a tighter bound on $|N_F(U) - S|$. In particular, we want to discard part of $S_a \cup S_z$.

Lemma 4.6 *There exists a subset $S' \subseteq S_a \cup S_z$ such that*

- (1) *S' has at most D^2 neighbors outside $S_a \cup S_z$, and*
- (2) *for some vertex v having a neighbor in S' , the set $S' \cup \{v\} \cup R_u$ is semi- k -distant from C , where $R_u = S_a - (S_a \cup S_z)$.*

Proof. Let $R_a = S_a - S_z$ and $R_z = S_z - S_a$. Suppose first that $|R_a| \geq |R_z|$. Whenever $x \in S_a$ and $y \in R_z$ in this proof, let $X_1 = N(x) \cap R_z$ and $X_2 = R_z - X_1$ and $S' = S_a \cup X_1$, and choose x^* that requires x and y' that requires y relative to S' . Define Q and Q' (they are disjoint) as in Lemma 2.5, with x_1 the neighbor of x on Q . This scenario is illustrated in Figure 6. We will show that when x is properly chosen, the resulting set S' satisfies conditions (1) and (2), often with $v = x_1$.

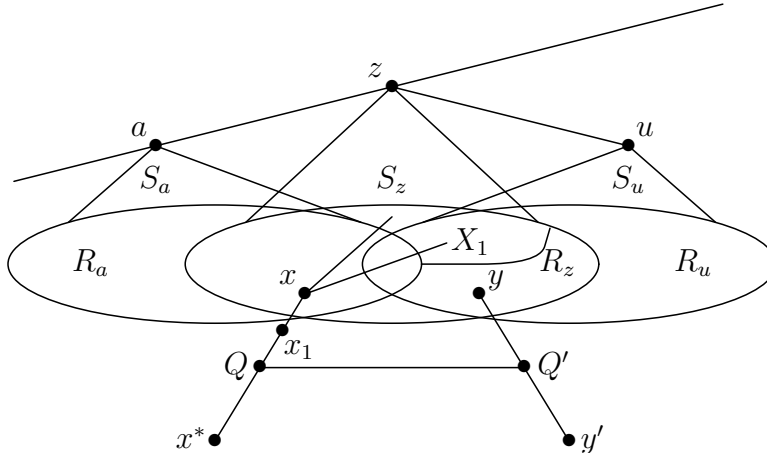


Figure 6: Setup for $S' = S_a \cup X_1$

Claim 1: If $xy \notin E(F)$, then $d_F(y', x_1) = k$ under either condition below:

(a) $x \in R_a$, or

(b) $x \in S_a - R_a$ and there is a vertex $v^* \in R_a \cap (N(y) - N(x))$.

In case (a), let P' be the subgraph induced by $\{x, a, z, y\}$. In case (b), let P' be the subgraph induced by $\{x, a, v^*, y\}$. In each case, P' is an induced x, y -path of length 3. By Lemma 4.3, this forbids $E(Q, Q') = \emptyset$. Hence $E(Q, Q') \neq \emptyset$, and Lemma 2.5(3) yields $d_F(y', x_1) = k$.

Now let $t = \min_{v \in R_a} |N(v) \cap R_z|$. Note that $t \leq |R_z|$. We consider two cases.

Case 1: Each vertex in $S_a - R_a$ has at least t neighbors in $S_a \cup S_z$. By the definition of t , in this case every vertex in S_a has at least t neighbors in $S_a \cup S_z$. Let x be a vertex of R_a with exactly t neighbors in R_z , so $|X_1| = t$. Consider any y' in C that requires a vertex y in X_2 relative to S' . Since $xy \notin E(F)$, Claim 1a yields $d(y', x_1) = k$. Now each vertex in C at distance k from X_2 has distance at most k from $S' \cup \{x_1\}$, so $S' \cup \{x_1\} \cup R_u$ is semi- k -distant from C . Finally, each vertex in S_a has at least t neighbors in $S_a \cup S_z$, and $|X_1| = t$, so

$$|N(S') - (S_a \cup S_z)| \leq |S_a|(D - t) + tD \leq D(D - t) + tD = D^2.$$

Case 2: Some vertex in $S_a - R_a$ has fewer than t neighbors in $S_a \cup S_z$. Among all such vertices in $S_a - R_a$, choose x to have the fewest neighbors in R_z , and let $t' = |X_1| = |N(x) \cap R_z|$. By the definitions of t and t' , each vertex in S_a has at least t' neighbors in $S_a \cup S_z$. Define a subset $X^* \subseteq X_2$ by putting y in X^* if y is required relative to S' by a vertex y' in C such that $d_F(y', x_1) > k$. We consider two subcases.

Subcase 2.1: $X^* = \emptyset$. This is the case where any y' in C that requires a vertex in X_2 relative to S' is at distance k from x_1 . Thus $S' \cup \{x_1\} \cup R_u$ is semi- k -distant from C , and

$$|N(S') - (S_a \cup S_z)| \leq |S_a|(D - t') + t'D \leq D(D - t') + t'D = D^2.$$

Subcase 2.2: $X^* \neq \emptyset$. If some vertex $v \in R_a$ has no neighbor in X^* , then we apply Claim 1a with v and v_1 in the role of x and x_1 . We conclude that each $y' \in C$ at distance k from X_2 has distance at most k from v_1 . Now $S' \cup \{v_1\} \cup R_u$ is semi- k -distant from C , and $|N(S') - (S_a \cup S_z)| \leq D^2$ by the computation in Subcase 2.1.

Hence we may assume that every $v \in R_a$ has a neighbor y in X^* . Since $d_F(y', x_1) > k$ for some vertex y' requiring y relative to S' , Claim 1b implies that $v \in N(x)$. Hence $R_a \subseteq N(x)$. Since $|R_a| \geq |R_z|$ and $|R_z| \geq t$ we obtain $|R_a| \geq t$, contradicting $|N(x) \cap (S_a \cup S_z)| < t$.

We have proved the lemma under the assumption that $|R_a| \geq |R_z|$. If instead $|R_z| \geq |R_a|$, then we repeat the entire argument of this lemma with the roles of a and z reversed. In steps where the original argument uses $P[a, w_p]$ to build an induced path of length $m - 1$, the reversed argument uses $P[z, w_{p-1}]$ when $x \in R_z$ and $P[a, w_p]$ when $x \in S_z \cap S_a$. ■

We now have all the tools for a final push to the theorem.

Theorem 4.7 *If $m > 5$ and m is even, then $\text{ex}^*(D; P_m) = \frac{1}{2}D^{m/2} + O(D^{m/2-1})$.*

Proof. Let $U = S' \cup \{v\} \cup R_u$, where $R_u = S_u - (S_a \cup S_z)$ and S' is the subset of $S_a \cup S_z$ guaranteed by Lemma 4.6. By Lemma 4.2, it suffices to prove that $|N_F(U) - S| \leq D^2 + O(D)$. We will discuss only neighborhoods within F and hence drop the subscript F . Since $|N(v)| \leq D$, it suffices to prove that $|N(S' \cup R_u) - S| \leq D^2$.

Let $R_{az} = S' - S_u$. By Lemma 4.4(2), each vertex of R_{az} is adjacent to each vertex of R_u . Lemma 4.5 implies that if $x \in S' \cap S_u$, then $N(x)$ contains R_u or R_{az} . Thus we partition $S' \cap S_u$ into R'_{az} and R'_u , where $R'_{az} = \{x \in S' \cap S_u: R_{az} \subseteq N(x)\}$ and $R'_u = (S' \cap S_u) - R'_{az}$ (see Figure 7). With $q = |R_{az}|$, $r = |R_u|$, $q' = |R'_{az}|$, and $r' = |R'_u|$, we have qr edges joining R_{az} and R_u , qq' edges joining R'_{az} and R_{az} , and rr' edges joining R'_u and R_u . Also, $S' \cup R_u = R_{az} \cup R'_{az} \cup R'_u \cup R_u$.

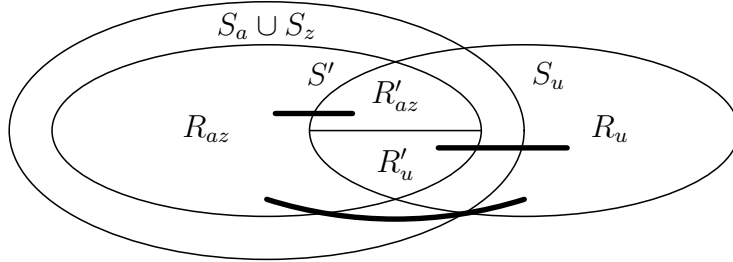


Figure 7: Forced edges in bold

When $q \leq r$, the proof is easy. Lower bounds on $|N(x) \cap S|$ are q for $x \in R'_{az} \cup R_u$ and r for $x \in R'_u$, where $r \geq q$. Therefore, $|N(x) - S| \leq D - q$ for $x \in R'_{az} \cup R_u \cup R'_u$. Since $|R_{az}| = q$, we obtain $|N(S' \cup R_u) - S| \leq (q' + r' + r)(D - q) + qD$. Since $R'_{az} \cup R'_u \cup R_u \subseteq S_u$, we have $q' + r' + r \leq D$. Hence $|N(S' \cup R_u) - S| \leq D(D - q) + qD = D^2$.

Therefore, we may henceforth assume that $q > r$. By Lemma 4.6, S' has at most D^2 neighbors outside $S_a \cup S_z$. More precisely, we computed in Lemma 4.6 that at most D^2 edges connect S' to $V(F) - (S_a \cup S_z)$. Since at least $r(q + r')$ of these go to R_u , and $R_u \subseteq S$, we conclude that $|N(S') - S| \leq D^2 - (q + r')r$.

On the other hand, we must allow for neighbors of R_u outside S . Each vertex of R_u is adjacent to $R_{az} \cup R'_u$ and hence has at most $D - (q + r')$ neighbors outside S . Thus,

$$|N(S' \cup R_u) - S| \leq D^2 - (q + r')r + r[D - (q + r')] = D^2 + r[D - 2(q + r')].$$

If $D < 2(q + r')$, then the proof is complete. Hence we may assume that $q + r' < D/2$.

Suppose first that $q' \leq D/2$. Now $|S'| = q' + q + r' \leq D/2 + D/2 = D$. Vertices of R'_{az} , R_{az} , and R'_u have at least q , $r + q'$, and r neighbors in S , respectively. Each lower bound is at least r , so each vertex of S' has at most $D - r$ neighbors outside S . Since each vertex of R_u has at most D neighbors outside S (using only that $r' + q \geq 0$), we have $|N(S' \cup R_u) - S| \leq |S'|(D - r) + |R_u|D \leq D(D - r) + rD = D^2$.

We may therefore assume that $q' > D/2$. Since $q + r' < D/2$, also $q' > r'$. We subtract as usual the incident edges forced within S for vertices of R_{az} , R'_{az} , R'_u , and R_u , respectively, to obtain our upper bound on $|N(S' \cup R_u) - S|$. We then regroup, add and subtract $r'q$, drop the nonpositive terms $-qr$ and $-2rr'$, and apply the inequalities $r' < q'$ and $q' + r' + r \leq D$ (since $q' + r' + r$ counts a subset of S_u , and $S_u \subseteq N(u)$). The resulting computation is

$$\begin{aligned} |N(S' \cup R_u) - S| &\leq q(D - q' - r) + q'(D - q) + r'(D - r) + r(D - q - r') \\ &= qD - q(q' + r) + (q' + r' + r)(D - q) + r'q - 2rr' \\ &\leq qD - qq' + D(D - q) + q'q = D^2. \quad \blacksquare \end{aligned}$$

By combining the ideas used in Theorems 3.2 and 4.7, it should be possible to show that also when m is even, $G_{D,m}$ achieves the maximum number of vertices among graphs in $\mathcal{F}(D, m)$. However, this would involve considerable additional detail. It would imply Theorem 4.7, but it would not improve the bound in Theorem 4.7, so we choose not to explore this detail.

Concerning D -regular graphs, we note that modifications to $G_{D,m}$ produce P_m -free D -regular graphs with more than $n(G_{D,m})/D^3$ vertices when D is odd (for $m \geq 8$) and more than $n(G_{D,m})/(D^3 2^{\lfloor m/2 \rfloor - 5})$ vertices when D is even (for $m \geq 11$).

5 Large P_5 -free graphs

In this section, we determine $\text{ex}^*(D, P_5)$. We first construct a family of graphs that contains the extremal graph. The construction yields $\text{ex}^*(D; P_5) \geq \frac{2}{27}D^3 + \frac{7}{18}D^2 + \frac{7}{18}D$.

Construction 5.1 Given a positive integer k , let $G_{D,k}$ denote the graph obtained from kK_{D-k+2} by adding edges to form a k -clique consisting of one vertex from each original component. The graph $G_{D,k}$ is P_5 -free and has $\binom{k}{2} + k\binom{D-k+2}{2}$ edges. As a function of k , the size is maximized when $k = \lceil (D+1)/3 \rceil$. The resulting number of edges is $\frac{2}{27}D^3 + \frac{7}{18}D^2 + c$, where c is $D/2$, $D/2 + 1/27$, or $7D/18 + 2/27$, depending on whether D is congruent to 0, 1, or 2 modulo 3.

We prove that when $D \geq 187$, $\text{ex}^*(D; P_5)$ has the form $\binom{k}{2} + k\binom{D-k+2}{2}$ for some k . Our approach relies on a structural property of P_5 -free graphs proved by Bacsó and Tuza [1].

Theorem 5.2 [1] *Every P_5 -free connected graph has a dominating set that is a clique or induces a 3-vertex path.*

We observed in the introduction that when D is even, $C_5[D/2]$ is in $\mathcal{F}(D; 2P_2)$ and has $5D^2/4$ edges; also $C_5[D/2] \in \mathcal{F}(D; P_5)$. When D is odd, expanding the vertices of C_5 into three independent sets of size $\lfloor D/2 \rfloor$ and two of size $\lceil D/2 \rceil$ yields a P_5 -free graph with maximum degree D and $\frac{1}{4}(5D^2 - 2D + 1)$ edges.

Proposition 5.3 *If G is a connected P_5 -free graph having maximum degree D and a dominating induced P_3 , then $e(G) \leq \frac{5}{4}D^2$. The bound is achieved by $C_5[D/2]$ when D is even.*

Proof. Let u, v, w in order be the vertices of the dominating P_3 ; note that $D \geq 2$. We show that $n(G) \leq \frac{5}{2}D$. If G has a dominating set of size at most 2, then $n(G) \leq 2D + 1 \leq \frac{5}{2}D$ (since G is connected and has maximum degree D). Hence we may assume that G has no dominating set of size at most 2. Let $S = \{u, v, w\}$.

Since S is a minimal dominating set in G , we have $N(x) \cap S = \{u\}$ for some $x \notin S$ and $N(y) \cap S = \{w\}$ for some $y \notin S$. Now x and y must be adjacent; otherwise, $\{x, u, v, w, y\}$ induces P_5 . Hence these vertices induce a 5-cycle C .

Every vertex outside C has a neighbor on C (in fact, in S). If some such vertex has exactly one neighbor on C , then there is an induced P_5 . Since at most $5(D-2)$ edges join $V(C)$ to the remaining vertices, we thus have $n(G) \leq 5 + 5(D-2)/2 = 5D/2$. ■

Theorem 5.4 *If $D \geq 187$, then $\text{ex}^*(D; P_5) = \max_k \left\{ \binom{k}{2} + k \binom{D-k+2}{2} \right\}$.*

Proof. Let G be a P_5 -free graph with maximum degree D . By Theorem 5.2, G has a dominating clique or P_3 . As observed in Construction 5.1, the claimed value is a lower bound and exceeds $5D^2/4$. By Proposition 5.3, we may therefore assume that G has a dominating clique W .

Let $k = |W|$. Since W is a dominating set, and each vertex in W has at most $D - k + 1$ neighbors outside W , we have $n(G) \leq k + k(D - k + 1) = k(D - k + 2)$ and $e(G) \leq k(D - k + 2)D/2$. If $k > 5D/6$ or $k < D/6$, then $n(G) < \frac{5}{36}D(D+2)$. Since $\frac{5}{36}D(D+2) < \frac{4}{27}D^2$ when $D > 30$, we may assume that $D/6 \leq k \leq 5D/6$.

Partition $V(G) - W$ into two sets A and B by the number of neighbors in W , letting $A = \{v \notin W: |N(v) \cap W| \geq 12\}$ and $B = \{v \notin W: |N(v) \cap W| \leq 11\}$ (Figure 8 illustrates this partition and the argument of the next paragraph). Let $a = |A|$ and $b = |B|$. Since at most $k(D - k + 1)$ edges leave W , we have $a \leq k(D - k + 1)/12$. If $b \leq 24D$, then $n(G) \leq k + k(D - k + 1)/12 + 24D < \frac{1}{4 \cdot 12}(D + 13)^2 + 24D$. Since $\frac{1}{4 \cdot 12}(D + 13)^2 + 24D < \frac{4}{27}D^2 + \frac{14}{18}D$ when $D \geq 187$, we may assume that $b > 24D$.

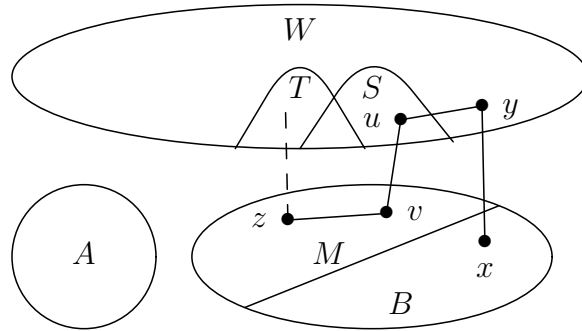


Figure 8: Forbidden P_5

We now claim that $\Delta(H) \leq D - k$, where H is the subgraph of G induced by B . If $|N(v) \cap B| > D - k$ for some $v \in B$, then choose $u \in N(v) \cap W$ (since W is a dominating

set, every vertex has a neighbor in W). Since u has at most $D - k$ neighbors other than v outside W , there exists $z \in (N(v) \cap B) - N(u)$. Let $S = N(v) \cap W$ and $T = N(z) \cap W$. The definition of B yields $|S \cap T| \leq 22$. Letting $M = N(S \cup T \cup \{v, z\})$, we thus have $|M| \leq 24D$. Since $b > 24D$, there exists $x \in B - M$. Choose $y \in N(x) \cap W$. Since $x \notin M$, we have $u, v, z \notin N(x)$ and $v, z \notin N(y)$; also z was chosen outside $N(u)$. Hence $\{x, y, u, v, z\}$ induces P_5 . This is a contradiction, so the claim holds.

The edges of G consist of those within W , those (at most $k(D - k + 1)$) from W to B , those in H , and the q edges incident to A . We have proved that $\Delta(H) \leq D - k$, so $e(H) \leq b(D - k)/2$. Since $|A| = a$, we have $q \leq aD$. Note that $aD \leq 12a(D - k)/2$, since $k \leq 5D/6$. Recall that $12a + b \leq k(D - k + 1)$. Summing the contributions now yields

$$e(G) \leq \binom{k}{2} + k(D - k + 1) + (12a + b)\frac{D - k}{2} \leq \binom{k}{2} + k\binom{D - k + 2}{2}. \quad \blacksquare$$

References

- [1] G. Bacsó, Zs. Tuza, Dominating cliques in P_5 -free graphs, *Period. Math. Hungar.* **21** (1990), 303–308.
- [2] F.R.K. Chung, A. Gyárfás, Zs. Tuza, W.T. Trotter, The maximum number of edges in a $2K_2$ -free graph, *Discrete Math.* **81** (1990). 129 – 135.
- [3] M.S. Chung, Thesis, University of Illinois, 1993.
- [4] M.S. Chung and D.B. West, Large P_4 -free graphs with bounded degree, *Journal of Graph Theory* **17** (1993), 109 – 116.
- [5] M.S. Chung and D.B. West, Large $2P_3$ -free graphs with bounded degree, in *Selected papers in honour of Paul Erdős on the occasion of his 80th birthday (Keszthely, 1993)*, *Discrete Math.* **150** (1996), 69–79.