Large P_4-Free Graphs with Bounded Degree

Myung S. Chung
Douglas B. West
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

ABSTRACT

Let $ex*(D;H)$ denote the maximum number of edges in a connected graph with maximum degree D and no induced subgraph isomorphic to H. We prove that this is finite only when H is a disjoint union of paths,m in which case we provide crude upper and lower bounds. When H is the four-vertex path P_4, we prove that the complete bipartite graph $K_{D,D}$ is the unique extremal graph. Furthermore, if G is a connected P_4-free graph with maximum degree D and clique number ω, then G has at most $D^2 - D(\omega - 2)/2$ edges. © 1993 John Wiley & Sons, Inc.

The archetypal problem of extremal graph theory is to determine the maximum number of edges in an n-vertex graph that does not contain some fixed graph H as a subgraph; this is usually written as $ex(n;H)$. Turán solved this when H is a clique; an extensive discussion of this and related problems appears in [1]. In this note, we study large graphs when we forbid H as an *induced* subgraph; such graphs are called *H-free*. Since any large clique is H-free if H is not a clique, and since disjoint copies of a graph introduce no new connected subgraphs, we define $ex^*(D;H)$ to be the maximum number of edges in a connected H-free graph G with maximum degree at most D. We solve the problem for $H = P_4$; our investigation is motivated by the solution of the problem for $H = 2K_2$ by Chung, Gyarfás, Trotter, and Tuza [3]. They proved that $ex^*(D;2K_2) = \frac{5}{4}D^2$ when D is even, attained uniquely by expanding each vertex of a 5-cycle into $D/2$ vertices with the same neighborhood (the bound is slightly smaller when D is odd). We prove that $ex^*(D;P_4) = D^2$, attained uniquely by $K_{D,D}$. In addition to a short proof that extends this bound to a larger class of graphs (those with disconnected complements and maximum degree D), we present a more detailed theorem that yields the maximum number of edges as a function

Journal of Graph Theory, Vol. 17, No. 1, 109–116 (1993)
 CCC 0364-9024/93/010109-08

also of the clique number $\omega(G)$, which is the maximum size of a vertex set inducing a complete subgraph of G. In this case the bound on the number of edges in G is $D^2 - D(\omega - 2)/2$.

Before presenting the results for P_4-free graphs, we consider the problem for general H. We prove that the value $ex^*(D;H)$ is finite only if H is a disjoint union of paths. Indeed, if H does not have this form, then there exist arbitrarily large H-free D-regular D-connected graphs. If H is a disjoint union of paths, then $ex^*(D;H) \leq \frac{1}{2}D^{m-1} + O(D^{m-2})$. For the m-vertex path P_m, we can at least show $ex^*(D,H) \geq \frac{1}{8}D^{\lceil (m+1)/2 \rceil}$.

1. THE GENERAL PROBLEM

Unboundness of $ex^*(D;H)$ is easy to obtain if H is not a disjoint union of paths. An arbitrarily long cycle will not contain any fixed cycle or any vertex of degree more than 2, so there are arbitrarily large 2-regular H-free connected graphs unless H is acyclic and has maximum degree 2. It is a bit more interesting to show that there are arbitrarily large H-free graphs with additional special properties. We use $n(G)$ and $e(G)$ to denote the number of vertices and edges of a graph G.

Theorem 1. If D is fixed and H is a fixed graph that is not a disjoint union of paths, then there are arbitrarily large H-free D-regular D-connected graphs.

Proof. We provide simple explicit constructions in several cases; many other constructions exist. If H contains a chordless cycle of length at least 4, then a simple generalization of the cycle mentioned above suffices. For any $N > Dn(H)/2$, we take $2N$ vertices on a cycle and form G by joining each vertex to the $\lfloor D/2 \rfloor$ closest vertices in each direction. If D is odd, we add a matching between opposite pairs of vertices on the cycle). The shortest chordless cycle of length at least 4 in G has about $4N/D$ vertices if D is even, about $2N/D$ if D is odd, so G is H-free. It is well known that the graphs constructed here are also D-connected (see [2]).

If H contains a triangle, suppose $N \geq 3$. Take $2N$ independent sets of size $\lfloor D/2 \rfloor$, indexed cyclically, and add all edges that join vertices from successive pairs of sets. If D is odd, add a matching between sets that are opposite on the cycle. The resulting graph is triangle-free, D-regular, and D-connected.

This leaves the case where H is a forest with a vertex of degree at least 3. Suppose $N \geq 1$. Take $2N$ cliques of size D. Partition the ith clique into vertex sets A_i and B_i of sizes $\lfloor D/2 \rfloor$, plus one leftover vertex x_i if D is odd. Add a matching between B_i and A_{i+1} (cyclically), and add the edge $x_i x_{i+N}$ if D is odd. The resulting graph G is D-regular and D-connected. Also, for any vertex, the subgraph induced by all but one of its neighbors is a clique; hence G is $K_{1,3}$-free and H-free. ∎ .

Thus the value $ex^*(D;H)$ can be finite only if H is a disjoint union of paths, and the next result provides a bound for such H.

Theorem 2. If H is a disjoint union of paths, then $ex^*(D;H)$ is finite. In particular, if $H = \sum_{i=1}^{k} P_{m_i}$ let $m = (\sum m_i) + k - 1$. Then $ex^*(D;H) \le \frac{1}{2}D^{m-1} + O(D^{m-2})$.

Proof. If G is H-free, then also G is P_m-free, since appropriate vertices of an induced P_m can be deleted to obtain H. Hence $ex^*(D;H) \le ex^*(D;P_m)$, and it suffices to consider $ex^*(D;P_m)$. A connected P_m-free graph has diameter less than $m - 1$. It is well known that a graph with maximum degree at most D and diameter less than $m - 1$ has at most $1 + D\sum_{i=1}^{m-2}(D - 1)^{i-1}$ vertices, because a vertex can fan out paths to at most $D(D - 1)^{i-1}$ vertices at each distance i. Since $e(G) \le Dn(G)/2$, we conclude that $ex^*(D;P_m)$ is bounded by a polynomial in D with leading term $\frac{1}{2}D^{m-1}$. ∎

This crude bound is within the square of the optimum for paths.

Construction of Large P_m-Free Graphs We construct for each $m \ge 3$ a P_m-free graph G_m having maximum degree D and at least a constant times $D^{\lceil m/2 \rceil}$ edges. First, $G_3 = K_{D+1}$ and $G_4 = K_{D,D}$. If $m = 5$, let G_m consist of a central clique Q with $\lfloor (D + 1)/3 \rfloor$ vertices and $\lfloor (D + 1)/3 \rfloor$ additional cliques of order $\lceil 2(D + 1)/3 \rceil$, each containing exactly one vertex of Q. The number of edges is $\frac{2}{27}D^3 + O(D^2)$.

If $m > 5$ and m is even, begin with the union of $D(D - 1)^{m/2-3}$ disjoint D-cliques. Form G_m by adding a tree with $m/2 - 1$ levels in which the root has D children, the other internal vertices have $D - 1$ children, and the leaves consist of one vertex from each of the cliques. The number of edges is $\frac{1}{2}D^{m/2} + O(D^{m/2-1})$.

If $m > 5$ and m is odd, the formation of G_m is more complex. Begin with the union of $\lfloor D^2/4 \rfloor (D - 1)^{(m-7)/2}$ disjoint D-cliques. Build $\lfloor D/2 \rfloor$ trees with $(m - 3)/2$ levels, in which the root has $\lceil D/2 \rceil$ children, the other internal nodes have $D - 1$ children, and the leaves consist of one vertex from each of the cliques. Complete G_m by adding an edge between every pair of roots. The number of edges is $\frac{1}{8}D^{(m+1)/2} + O(D^{(m-1)/2})$.

Other Constructions. Suppose $m \ge 5$ and m is even. Then the graph G_{m+2} shows that $ex^*(D;P_m + P_{m-2}) \ge \frac{1}{2}D^{m/2+1} + O(D^{m/2})$, because any induced copy of P_m must contain the root, and any induced copy of P_{m-2} must contain a neighbor of the root. For $m \ge 5$ and m odd, the graph G_{m+2} shows that $ex^*(D;2P_m) \ge \frac{1}{8}D^{\lceil m/2 \rceil+1} + O(D^{\lceil m/2 \rceil})$, because any induced copy of P_m must contain a vertex of the central clique. For small m, we can make stronger statements: G_8 shows that $ex^*(D;P_6 + P_3) \geq \frac{1}{2}D^4 + O(D^3)$, and G_7 shows that $ex^*(D;2P_3) \geq \frac{1}{8}D^4 + O(D^3)$. We have now proved that G_7 is the largest $2P_3$-free graph, and we will present this result in a subsequent paper.

For more general disjoint unions of paths, it is more difficult to construct large H-free graphs. When H contains at least two paths, a $2K_2$-free graph is also automatically H-free, which implies $ex^*(D;H) \geq \frac{5}{4}D^2$. We can improve the leading constant as follows. If $H = \sum_{i=1}^{k} P_{m_i}$, let $m = (\sum m_i) + k - 1$. If D is even, then the graph obtained by expanding each vertex of an m-cycle into $D/2$ vertices with the same neighborhood is H-free and has $(m/4)D^2$ edges. At present, we have no examples of connected kP_2-free graphs with more edges than this.

2. DETERMINATION OF $ex^*(D;P_4)$

In this section we provide a short proof that $ex^*(D;P_4) = D^2$. This approach was suggested by Stephan Olariu, who independently found another proof of the second result here. This approach is based on a well-known property of P_4-free graphs, which is that a connected P_4-free graph has a disconnected complement. We then obtain the bound of D^2 on the size of connected P_4-free graphs with maximum degree D by proving the more general result that this bound holds also for graphs with maximum degree D that have a disconnected complement, and here again the unique extremal graph is $K_{D,D}$.

We use the following notation. The set of vertices adjacent to x is $N(x)$, and $x \leftrightarrow y$ means "x is adjacent to y"; i.e., $\leftrightarrow$ denotes the adjacency relation on the vertices. We extend the notation for adjacency to sets of vertices, writing $S \leftrightarrow T$ if every member of S is adjacent to every member of T. The *join* of two graphs G_1, G_2 is the graph $G = G_1 \vee G_2$ obtained by taking the disjoint union $G_1 + G_2$ and adding edges so that $V(G_1) \leftrightarrow V(G_2)$.

Theorem 3. If $\Delta(G) \leq D$ and $\overline{G}$ is disconnected, then $e(G) \leq D^2$, with equality only for $K_{D,D}$.

Proof. Since $\overline{G}$ is disconnected, G is the join of two subgraphs G_1, G_2; let $n_i = n(G_i)$. The degree bound implies $n_i \leq D$. There are several ways to complete the proof.

Proof 1. If $n_i = D$, then the degree bound implies G_{3-i} is an independent set. Hence if $G \neq K(D,D)$, we may assume $n(G_1) < D$. We form a new graph G' with more edges. Let H be a spanning forest of G_2. Form G' by deleting the edges of H and adding a new vertex x to G_1 that is adjacent to every vertex of G_2. We have deleted at most $n(G_2) - 1$ edges and added $n(G_2)$ edges, so $e(G') > e(G)$. For every nonisolated vertex of G_2, we have deleted at least one edge (in H) and added an incident edge to x, so the degree bound still holds. For every isolated vertex of G_2, we have added one neighbor, but $n(G_1) < D$ guarantees that the degree bound still holds. Finally, the complement of G' is still disconnected, as it has no edge from $V(G_1) \cup \{x\}$ to $V(G_2)$.

Proof 2 (provided by Stephen Seng-Wah Kwek). From $G = G_1 \vee G_2$, we conclude $\Delta(G_i) \leq D - n_{3-i}$. Hence $e(G) \leq n_1 n_2 + n_1(D - n_2)/2 + n_2(D - n_1)/2 = (n_1 + n_2)D/2 \leq D^2$. Furthermore, equality requires $n_1 + n_2 = 2D$, hence $n_1 = n_2 = D$, which as observed above forces $G = K_{D,D}$.

Proof 3 (provided by Sannjeev Khanna and Jalal Wehbeh). If $\Delta(G) \leq D$, then $\delta(\overline{G}) \geq n(G) - 1 - D$. If we choose u, v from different components of $\overline{G}$, we have $2(n(G) - 1 - D) \leq d_{\overline{G}}(u) + d_{\overline{G}}(v) \leq n(G) - 2$, or $n(G) \leq 2D$. Hence $e(G) \leq nD/2 \leq D^2$. Equality requires $n(G) = 2D$ and $\delta(\overline{G}) = D - 1$, so $\overline{G}$ must be $2K_D$.

Corollary 1. $ex^*(D; P_4) = D^2$, achieved uniquely by $K_{D,D}$.

Proof. This follows immediately from the fact that a connected P_4-free graph has a disconnected complement, proved by Seinsche [4]. For completeness, we include here a short proof by induction on $n(G)$. As a basis, the complement of any connected 3-vertex graph is disconnected. For $n(G) > 3$, choose a spanning tree with x a leaf and u its neighbor, so $G - x$ is connected. Since $G - x$ is also P_4-free, the induction hypothesis guarantees a partition A, B of $V(G) - x$ such that $A \leftrightarrow B$; we may assume $u \in A$. If $\overline{G}$ is not disconnected, then x has a non-neighbor in A and in B. Choose $z \in B - N(X)$, and let $A' = A \cap N(x)$. For any $y \in A'$ and $w \in A - A'$, we have x, y, z, w inducing P_4 unless $y \leftrightarrow w$. Hence $A' \leftrightarrow A - A'$, but now $A' \leftrightarrow V(G) - A'$ and $\overline{G}$ is disconnected. ∎

3. P_4-FREE GRAPHS WITH FIXED CLIQUE NUMBER

In addition to $\leftrightarrow$ for adjacency, we use $\not\leftrightarrow$ for nonadjacency. We view a clique as a set of vertices inducing a complete subgraph, so when we say Q is a clique, we write $v \in Q$ to indicate a vertex of Q. A *dominating clique* in G is a clique containing a neighbor of every vertex of G.

Lemma 1. In a connected P_4-free graph, any maximal clique is a dominating clique.

Proof. If Q is a nondominating maximal clique, then there is a vertex x with no neighbor in Q that has a common neighbor y with a vertex

$z \in Q$. Since Q is maximal, there exists $w \in Q$ such that $y \not\leftrightarrow w$. But then $\{x, y, z, w\}$ induces P_4. ∎

Henceforth, we fix Q to be a particular maximum clique, with size ω. Given $A \subseteq Q$, we define $S_A = \{v \in V(G): N(v) \cap Q = A\}$. By Lemma 1, the sets S_A for nonempty A partition $V(G)$.

Lemma 2. If $S_A, S_{A'}$ are nonempty and A, A' are not ordered by inclusion, then $S_A \leftrightarrow S_{A'}$ and $A \cup A' = Q$. ∎

Proof. Consider $v \in A - A'$ and $v' \in A' - A$, and select $x \in S_A$ and $x' \in S_{A'}$. The vertices $\{x, v, v', x'\}$ induce P_4 unless $x \leftrightarrow x'$. Since $x \leftrightarrow x'$, the existence of any $u \in Q - (A \cup A')$ produces a P_4 induced by $\{x, x', v', u\}$.

Lemma 3. If $A = Q - u$, then S_A is a nonempty stable set, and $u \in A'$ implies $S_A \leftrightarrow S_{A'}$. In particular, $\bigcup_{v \in Q} S_{Q-v}$ induces a complete ω-partite graph Q' containing Q, with partite sets $\{S_{Q-v}\}$.

Proof. The set S_A is nonempty because it contains u. If there is an edge within S_A, then G has a larger clique. The claim for A' is a special case of Lemma 2. The second sentence follows immediately from the first. ∎

Lemma 4. The inclusion ordering P on $A \subseteq Q : S_A \neq \emptyset$ is a rooted forest, meaning that no element of P covers two distinct elements. The maximal elements of P are $\{Q - v : v \in Q\}$, and every element of Q is omitted by exactly one minimal element of P.

Proof. Since Q is a maximum clique, S_Q is empty; also all S_{Q-v} are nonempty. If any set A in P contained two incomparable sets B, C in P, then $B \cup C \subseteq A \subset Q$, contradicting Lemma 2. Because the maximal elements are $\{Q - v\}$, every element of Q is omitted from at least one minimal element. Two minimal elements B, C omitting v would be incomparable sets contradicting Lemma 2. ∎

Theorem 4. If G is a P_4-free connected graph with maximum degree D and clique number ω, then $e(G) \le D^2 - D(\omega - 2)/2$.

Proof. It suffices to prove $n(G) \le 2D - \omega + 2$; we illustrate the argument by first proving $n(G) \le 2D$. Fix the maximum clique Q and resulting poset P as described earlier. Let $A_1, \ldots, A_t$ be the minimal sets in P, and let $S_i = \bigcup_{A \supseteq A_i} S_A$. The sets S_i partition $V(G)$. Note that $t \ge 2$, because Q is dominating and no element appears in all sets $Q - v$. Select an element $v_i \in Q - A_i$ for each i. By Lemma 4, v_i appears in each other minimal set in P, and then Lemma 2 implies $v_i \leftrightarrow V(G) - S_i$. Hence

$n(G) - |S_i| \leq D$. Summing this over i yields $(t-1)n(G) \leq tD$. Since $t \geq 2$, we have $n(G) \leq 2D$.

We refine this argument to obtain $n(G) \leq 2D - \omega + 2$. Suppose first that some maximal element $Q - v$ of P is a minimal element; i.e., $A_j = Q - v$ and $S_j = S_{Q-v}$. As above, $v \leftrightarrow V(G) - S_j$. For any $u \in Q$ with $u \neq v$, we have $u \in S_{Q-u}$ and hence $u \leftrightarrow S_j \bigcup (Q - u)$; this implies $|S_j| \leq d(u) - \omega + 2$, because $S_j \cap (Q - u) = \{v\}$. Together with $n(G) - |S_j| \leq d(v)$, this implies $n(G) \leq 2D - \omega + 2$.

When all the sets $Q - v$ are nonminimal in P, we define $P' = P - \{Q - v\}$ to be the nonmaximal sets in P, and let $S_i' = \bigcup_{A_i \subseteq A \in P'} S_A$. For $v \in Q$ and $A \subseteq Q$, define $m(v) = |S_{Q-v}|$, and $m(A) = \sum_{v \in A} m(v)$. The vertices not in $\bigcup S_i'$ are the vertices of the complete multipartite graph Q' that Lemma 3 implies is induced by $\bigcup S_{Q-v}$; hence $n(G) = m(Q) + \sum |S_i'|$. Again selecting $v_i \in Q - A_i$, we have $v_i \leftrightarrow V(G) - S_{Q-v_i} - S_i'$, or $n(G) - m(v_i) - |S_i'| \leq D$. Summing this over i and applying $\sum |S_i'| = n(G) - m(Q)$, we have $(t-1)n(G) + m(Q) - \sum_{i=1}^{t} m(v_i) \leq tD$. Since the summation omits $\omega - t$ partite sets in Q', the sum differs from $m(Q)$ by at least $\omega - t$. We conclude that $n(G) \leq (tD + t - \omega)/(t-1) \leq 2(D+1) - \omega$, the latter inequality holding for $t \geq 2$. ∎

By Lemma 1, every P_4-free graph with $\omega = 2$ is a complete bipartite graph. Hence the bound proved above also implies the uniqueness of the extremal graph $K_{D,D}$ in achieving $ex * (D; P_4)$.

It is interesting to note that in each of the two cases of the proof of $n(G) \leq 2D - \omega + 2$ above, it is possible to achieve equality throughout and construct D-regular P_4-free connected graphs with $2D - \omega + 2$ vertices. We close by describing these constructions.

Theorem 5. For each $\omega > 2$, the bound in theorem 4 is best possible for infinitely many D, but the extremal graph is not generally unique.

Proof. The first type of construction can be used to achieve the bound when $D - \omega + 1$ is a multiple of $\omega - 1$ or even multiple of $\omega - 2$, and in other settings. The sets in P will consist of the maximal sets $\{Q - u : u \in Q\}$ and a single nonmaximal set that is a singleton $\{v\}$. In this case $Q - v$ is both maximal and minimal, and $\{Q - v, \{v\}\}$ are the only minimal sets; call them A_1, A_2, respectively. Then $S_1 = S_{Q-v}$. We want $|S_1| = D - \omega + 2$, so choose $Q' = K_{D-\omega+2,1,\ldots,1}$ (altogether $D + 1$ vertices). The remaining $D - \omega + 1$ vertices will all be adjacent to all the $D - \omega + 2$ vertices of S_{Q-v}. It suffices to let $G - Q'$ be a P_4-free $(\omega - 2)$-regular graph, if such a graph exists. Of course, this requires $\omega \leq (D+2)/2$ and other parity or divisibility conditions. If $\omega = 2$, the resulting graph is always $K_{D,D}$. If $\omega - 1$ divides $D - \omega + 1$, then we can take $G - Q'$ to be disjoint $(\omega - 1)$-cliques. If $D - \omega + 1$ is an even multiple of $\omega - 2$, then we can take $G - Q'$ to be disjoint copies of

$K_{\omega-2,\omega-2}$. Since $\omega - 1$ and $\omega - 2$ are relatively prime, we can combine these constructions to achieve the bound whenever $D - \omega + 1$ is even and D is sufficiently large.

Another construction also depends on the existence of suitable P_4-free regular graphs, which is easy to verify for many large values of D. As in the second case in Theorem 4, all maximal elements $Q - v$ of P are nonminimal, and we have only two other sets A, A' in P. Here we take $Q' = Q = K_\omega$, and $G - Q$ consists of the two sets S_A and $S_{A'}$, with $|S_A| = |S_{A'}| = D - \omega + 1$. By Lemma 2, we must have $S_A \leftrightarrow S_{A'}$ and $A \cup A' = Q$ in such a construction. Let $A' = Q - A$, with $k = |A|$. The only unspecified edges are within the subsets induced by S_A and $S_{A'}$. We complete a P_4-free D-regular graph with clique number ω if the subgraph induced by S_A is P_4-free and $(\omega - 1 - k)$-regular and the subgraph induced by $S_{A'}$ is P_4-free and $(k - 1)$-regular. When D is large it is often possible to construct two regular P_4-free graphs on $D - \omega + 1$ vertices having degrees $\omega - 1 - k$ and $k - 1$. ∎

ACKNOWLEDGMENT

This research was supported in part by ONR Grant N00014-85K0570 and by the University of Illinois Research Board.

References

[1] B. Bollobás, *Extremal Graph Theory*, (Academic Press, 1978) 292–367.
[2] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, (Elsevier North-Holland, 1976) 48–49.
[3] F. R. K. Chung, A. Gyárfás, Z. Tuza, and W. T. Trotter, The maximum number of edges in a $2K_2$-free graph. *Discrete Math.* **81** (1990) 129–135.
[4] D. Seinsche, On a property of the class of n-colorable graphs. *J. Combinat. Theory B* **16** (1974) 191–193.