Graphs with many copies of a given subgraph

Vladimir Nikiforov

Department of Mathematical Sciences,
University of Memphis, Memphis TN 38152

vnikifr@memphis.edu

Submitted: Oct 8, 2007; Accepted: Mar 3, 2008; Published: Mar 12, 2008
Mathematics Subject Classification: 05C35

Abstract

Let $c > 0$, and H be a fixed graph of order r. Every graph on n vertices containing at least cn^r copies of H contains a "blow-up" of H with $r-1$ vertex classes of size $\lfloor c^{r^2} \ln n \rfloor$ and one vertex class of size greater than $n^{1-c^{r-1}}$. A similar result holds for induced copies of H.

Main results

Suppose that a graph G of order n contains cn^r copies of a given subgraph H on r vertices. How large "blow-up" of H must G contain? When H is an r-clique, this question was answered in [3]: G contains a complete r-partite graph with $r-1$ parts of size $\lfloor c^r \ln n \rfloor$ and one part larger than $n^{1-c^{r-1}}$.

The aim of this note is to answer this question for any subgraph H.

First we define precisely a "blow-up" of a graph: given a graph H of order r and positive integers $x_1, \ldots, x_r$, we write $H(x_1, \ldots, x_r)$ for the graph obtained by replacing each vertex $u \in V(H)$ with a set V_u of size x_u and each edge $uv \in E(H)$ with a complete bipartite graph with vertex classes V_u and V_v.

Theorem 1 *Let $2 \leq r \leq n$, $(\ln n)^{-1/r^2} \leq c \leq 1/4$, H be a graph of order r, and G be a graph of order n. If G contains more than cn^r copies of H, then G contains a copy of $H(s, \ldots s, t)$, where $s = \lfloor c^{r^2} \ln n \rfloor$ and $t > n^{1-c^{r-1}}$.*

To state a similar theorem for induced subgraphs, we need a proper modification of $H(x_1, \ldots, x_r)$: we say that a graph X *is of type* $H(x_1, \ldots, x_r)$, if X is obtained from $H(x_1, \ldots, x_r)$ by adding some (possibly zero) edges within the sets V_u, $u \in V(H)$.

Theorem 2 *Let $2 \leq r \leq n$, $(\ln n)^{-1/r^2} \leq c \leq 1/4$, H be a graph of order r, and G be a graph of order n. If G contains more than cn^r induced copies of H, then G contains an induced subgraph of type $H(s, \ldots s, t)$, where $s = \lfloor c^{r^2} \ln n \rfloor$ and $t > n^{1-c^{r-1}}$.*

Remarks

- The relations between c and n in Theorems 1 and 2 need some explanation. First, for fixed c, they show how large must be n to get valid conclusions. But, in fact, the relations are subtler, for c itself may depend on n, e.g., letting $c = 1/\ln \ln n$, the conclusions are meaningful for sufficiently large n.

- Note that, in Theorems 1 and 2, if the conclusion holds for some c, it holds also for $0 < c' < c$, provided n is sufficiently large.

- The exponent $1 - c^{r-1}$ in Theorems 1 and 2 is not the best one, but is simple.

- Using random graphs, it is easy to see that most graphs on n vertices contain substantially many copies of any fixed graph, but contain no complete bipartite subgraphs with both parts larger than $C \log n$, for some $C > 0$, independent of n. Hence, Theorems 1 and 2 are essentially best possible.

General notation

Our notation follows [1]; thus, given a graph G, we write:

- $V(G)$ for the vertex set of G;
- $E(G)$ for the edge set of G and $e(G)$ for $|E(G)|$;
- K_2 for the complete graph of order 2;
- $K_2(s,t)$ for the complete bipartite graph with parts of size s and t;
- $f|_X$ for the restriction of a map f to a set X.

Specific notation

Suppose that G and H are graphs, and let X be an induced subgraph of H.

- We write $H(G)$ for the set of injections $h : V(H) \to V(G)$, such that $\{u,v\} \in E(H)$ if and only if $\{h(u), h(v)\} \in E(G)$.

- We say that $h \in H(G)$ *extends* $g \in X(G)$, if $g = h|_{V(X)}$.

Suppose that $M \subset H(G)$.

- We let

$$X(M) = \{g : (g \in X(G)) \ \& \ (\text{there exists } h \in M \text{ extending } g)\}.$$

- For every $g \in X(M)$, we let

$$d_M(g) = \left| \{h : (h \in M) \ \& \ (h \text{ extends } g)\} \right|.$$

Suppose that Y is a subgraph of G of type $H(s_1, \ldots, s_r)$ and let $s = \min\{s_1, \ldots, s_r\}$.

- We say that M *covers* Y if:

  (a) for every edge ij going across vertex classes of Y, there exists $h \in M$ mapping some edge of H onto ij;

  (b) there exists $h_1, \ldots, h_s \in M$, such that $h_i(H) \cap h_j(H) = \varnothing$ for $i \neq j$, and for all $i \in [s]$, $h_i(H)$ contains a vertex from each vertex class of Y.

Condition (b) implies that if M covers Y, then Y contains s disjoint images of H, which are mapped via injections from M and which contain exactly one vertex from each vertex class of Y. This technicality is needed for a proof by induction.

Proofs

The proofs of Theorems 1 and 2 are almost identical, so we shall present only the proof of Theorem 2, for it needs more care. We deduce Theorem 2 from the following technical statement.

Theorem 3 *Let $2 \leq r \leq n$, $(\ln n)^{-1/r^2} \leq c \leq 1/4$, H be a graph of order r, and G be a graph of order n. If $M \subset H(G)$ and $|M| \geq cn^r$, then M covers an induced subgraph of type $H(s, \ldots s, t)$ with $s = \lfloor c^r 4^{-r^2+r} \ln n \rfloor$ and $t > n^{1-c^{r-1}}$.*

To see that Theorem 3 implies Theorem 2, note that to each induced copy of $H \subset G$ corresponds an injection $h \in H(G)$, and to different copies correspond different injections. Hence, if G contains cn^r induced copies of H, we have a set $M \subset H(G)$ with $|M| \geq cn^r$. By Theorem 3, G contains an induced subgraph Y of type $H(s, \ldots s, t)$ with $s = \lfloor c^r 4^{-r^2+r} \ln n \rfloor$ and $t > n^{1-c^{r-1}}$; now Theorem 2 follows, in view of $c^r 4^{-r^2+r} \geq c^{r^2}$.

In turn, the proof of Theorem 3 is based on the following lemma.

Lemma 4 *Let F be a bipartite graph with parts A and B. Let $2 \leq r \leq n$, $(\ln n)^{-1/r^2} \leq c \leq 1/2$, $|A| = m$, $|B| = n$, and $s = \lfloor c^r 4^{-r^2+r} \ln n \rfloor$. If $s \leq (c/2^r)m + 1$ and $e(F) \geq (c/2^{r-1})mn$, then F contains a $K_2(s, t)$ with parts $S \subset A$ and $T \subset B$ such that $|S| = s$ and $|T| = t > n^{1-c^{r-1}}$.*

Proof Let

$$t = \max \{x : \text{there exists } K_2(s, x) \subset F \text{ with part of size } s \text{ in } A\}.$$

For any $X \subset A$, write $d(X)$ for the number of vertices joined to all vertices of X. By definition, $d(X) \leq t$ for each $X \subset A$ with $|X| = s$; hence,

$$t\binom{m}{s} \geq \sum_{X \subset A, |X|=s} d(X) = \sum_{u \in B} \binom{d(u)}{s}. \tag{1}$$

Following [2], p. 398, set

$$f(x)=\begin{cases}\binom{x}{s} & \text{if } x\geq s-1\\ 0 & \text{if } x<s-1,\end{cases}$$

and note that $f(x)$ is a convex function. Therefore,

$$\sum_{u\in B}\binom{d(u)}{s}=\sum_{u\in B}f\left(d(u)\right)\geq nf\left(\frac{1}{n}\sum_{u\in B}d(u)\right)=n\binom{e(F)/n}{s}\geq n\binom{cm/2^{r-1}}{s}.$$

Combining this inequality with (1), and rearranging, we find that

$$\begin{aligned}t&\geq n\frac{(cm/2^{r-1})(cm/2^{r-1}-1)\cdots(cm/2^{r-1}-s+1)}{m(m-1)\cdots(m-s+1)}>n\left(\frac{cm/2^{r-1}-s+1}{m}\right)^s\\&\geq n\left(\frac{c}{2^r}\right)^s\geq n\left(e^{\ln(c/2^r)}\right)^{c^r4^{-r^2+r}\ln n}=n^{1+c^r4^{-r^2+r}\ln(c/2^r)}.\end{aligned}$$

Since $c/2^r\leq 1/8<1/e$ and $x\ln x$ is decreasing for $0<x<1/e$, and in view of

$$\frac{2^{2r^2-2r-1}}{r+1}\geq 1\geq \ln 2,$$

we see that

$$\begin{aligned}c4^{-r^2+r}\ln(c/2^r)&\geq\left((c/2^r)\ln(c/2^r)\right)2^{-2r^2+3r}\geq-\left(2^{-r+1}\left(r+1\right)\right)2^{-2r^2+3r}\\&\geq-\left(r+1\right)2^{-2r^2+2r+1}\ln 2\geq -1.\end{aligned}$$

Now, $c^r4^{-r^2+r}\ln\left(c/2^r\right)\geq -c^{r-1}$ and so,

$$t>n^{1+c^r4^{-r^2+r}\ln(c/2^r)}\geq n^{1-c^{r-1}}.$$

$\square$

Proof of Theorem 3 Let $M\subset H(G)$ satisfy $|M|\geq cn^r$. To prove that M covers an induced subgraph of type $H(s,\ldots s,t)$ with $s=\left\lfloor c^r4^{-r^2+r}\ln n\right\rfloor$ and $t>n^{1-c^{r-1}}$ we shall use induction on r.

Assume $r=2$ and let A and B be two disjoint copies of $V(G)$. We can suppose that $H=K_2$, as otherwise we can apply the subsequent argument to the complement of G.

Let us define a bipartite graph F with parts A and B, joining $u\in A$ to $v\in B$ if $uv\in M$. Set $s=\left\lfloor(c^2/16)\ln n\right\rfloor$ and note that $s\leq(c/4)n+1$. Since $e(F)=|M|\geq cn^2>(c/2)n^2$, Lemma 4 implies that F contains a $K_2(s,t)$ with $t>n^{1-c}$. Hence M covers an induced graph of type $K_2(s,t)$, proving the assertion for $r=2$. Now let $r>2$ and assume the assertion true for $r-1$.

Let $V(H)=\{v_1,\ldots,v_r\}$ and $H'=H\left[\{v_1,\ldots,v_{r-1}\}\right]$.

We first show that there exists $L\subset M$ with $|L|>(c/2)n^r$ such that $d_L(h)>(c/2)n$ for all $h\in H'(L)$. Indeed, set $L=M$ and apply the following procedure.

While *there exists an* $h \in H'(L)$ *with* $d_L(h) \leq (c/2)n$ **do**

Remove from L *all members extending* h.

When this procedure stops, we have $d_L(h) > (c/2)n$ for all $h \in H'(L)$, and also

$$|M| - |L| \leq \frac{c}{2}n\,|H'(M)| < \frac{c}{2}n \cdot n^{r-1},$$

giving $|L| > (c/2)n^r$, as claimed.

Since $H'(L) \subset H'(G)$ and

$$|H'(L)| \geq |L|/n > (c/2)n^r/n = (c/2)n^{r-1},$$

the induction assumption implies that $H'(L)$ covers an induced subgraph $R \subset G$ of type $H'(p,\ldots,p)$ with $p = \left\lfloor (c/2)^{r-1}4^{-(r-1)^2+r-1}\ln n \right\rfloor$. Here we use the inequalities

$$n^{1-c^{r-2}} \geq n^{1-c} \geq n^{1/2} > 2^{-4}\ln n \geq (c/2)^{r-1}4^{-(r-1)^2+r-1}\ln n.$$

Write $U_1,\ldots,U_{r-1}$ for the vertex classes of R. Since $H'(L)$ covers R, we know that there exist $h_1,\ldots,h_p \in H'(L)$ such that $h_1(H'),\ldots,h_p(H')$ are disjoint subgraphs of R containing a vertex from U_i, for all $i \in [r-1]$. For every $i \in [p]$, let

$$W_i = \big\{v : (\text{there exists } g \in L \text{ extending } h_i) \ \& \ (g(v_r) = v)\big\}.$$

That is to say, each vertex in W_i together with the vertices of $h_i(H')$ induces a copy of H.

Write d for the degree of v_r in H and note that each $v \in W_i$ is joined to exactly d vertices of $h_i(H')$. Since by our selection, $|W_i| = d_L(h_i) \geq (c/2)n$ for all $i \in [p]$, there is a set $X_i \subset W_i$ with

$$|X_i| \geq (cn/2)/\binom{r-1}{d} \geq \frac{cn}{2^{r-1}}$$

such that all vertices of X_i have the same d neighbors in $h_i(H')$. Let $Y_i \subset [r-1]$ be defined as

$$Y_i = \{j : U_j \text{ contains a neighbor of a vertex in } X_i\}.$$

Each of the sets $Y_1,\ldots,Y_p$ is a d-element subset of $[r-1]$; by the pigeonhole principle, there exists a set $A \subset [p]$ with

$$|A| \geq p/\binom{r-1}{d} \geq \left\lceil p/2^{r-2} \right\rceil$$

such that the sets Y_i are the same for all $i \in A$. Note that for every $i \in A$ and every $v \in X_i$, the neighbors of v in $h_i(H')$ belong exactly to the same d vertex classes of R. Letting $m = \lceil p/2^{r-2} \rceil$, we may and shall assume that $|A| = m$.

Let us define a bipartite graph F with parts A and $B = V(G)$, joining $i \in A$ to $v \in B$ if $v \in X_i$. Since $|X_i| > cn/2^{r-1}$ for all $i \in A$, we see that

$$e(F) > \frac{c}{2^{r-1}}mn.$$

Also, setting $s = \left\lfloor c^r 4^{-r^2+r} \ln n \right\rfloor$, we find that

$$\begin{aligned} s &\leq c^r 4^{-r^2+r} \ln n = (c2^{-3r-3}) \left((c/2)^{r-1} 4^{-(r-1)^2+r-1} \right) \ln n \\ &< (c2^{-2r-2}) \left\lfloor (c/2)^{r-1} 4^{-(r-1)^2+r-1} \ln n \right\rfloor + 1 \leq (c/2^r)(p/2^{r-2}) + 1 \\ &\leq (c/2^r)m + 1. \end{aligned}$$

By Lemma 4, F contains a complete bipartite graph $K_2(s,t)$ with parts $S \subset A$ and $T \subset B = V(G)$ such that $|S| = s$ and $|T| = t > n^{1-c^{r-1}}$.

Let $G' = G\left[\cup_{i\in S} h_i(H')\right]$ and $G'' = G\left[\cup_{i\in S} h_i(H') \cup T\right]$. Note that G' is an induced subgraph of R and so G' is of type $H'(s,\ldots,s)$. To prove that G'' is of type $H(s,\ldots,s,t)$ select $v \in T$ and $h \in L$ such that $h|_{V(H')} = h_1$ and $h(v_r) = v$. By our construction v has exactly d neighbors in $h_1(H')$, belonging say to the vertex classes $U_1,\ldots,U_d$. Since all neighbors of v in G' belong to the same vertex classes, and v has d neighbors in each $h_2(H'),\ldots,h_s(H')$, we see that v is joined to every vertex in $\cup_{i=1}^d U_i$, and is not joined to any vertex in $V(G')\backslash(\cup_{i=1}^d U_i)$. Since this holds for all vertices $v \in T$, we see that G'' is of type $H(s,\ldots,s,t)$.

To finish the proof, we shall show that L covers G''. By the induction assumption, L covers R, hence for every edge ij going across vertex classes of G', there exists $h \in L$ mapping some edge of H onto ij. On the other hand, let $u \in h_i(H')$ be joined to $v \in T$; by our construction there exist $h \in L$ such that $h|_{V(H')} = h_i$ and $h(v_r) = v$. Thus, $h^{-1}(u)v \in E(H)$, and h maps an edge of H onto uv. This proves condition (a) for covering.

Finally, taking s distinct vertices $u_1,\ldots,u_s \in T$, by the construction of T, for every $i \in S$, there exists $g_i \in L$ with $g_i|_{V(H')} = h_i$ and $g_i(v_r) = u_i$. Hence, L covers G'', completing the induction step and the proof of Theorem 3. $\square$

References

[1] B. Bollobás, *Modern Graph Theory,* Graduate Texts in Mathematics, 184, Springer-Verlag, New York (1998).

[2] L. Lovász, *Combinatorial problems and exercises*, North-Holland Publishing Co., Amsterdam-New York (1979).

[3] V. Nikiforov, Graphs with many r-cliques have large complete r-partite subgraphs, to appear in *Bull. London Math. Soc.*