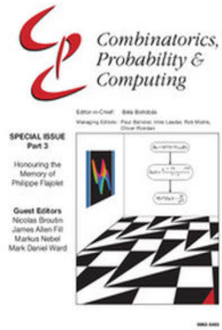

Maximizing the Number of Independent Sets of a Fixed Size

WENYING GAN, PO-SHEN LOH and BENNY SUDAKOV

Combinatorics, Probability and Computing / Volume 24 / Issue 03 / May 2015, pp 521 - 527
DOI: 10.1017/S0963548314000546, Published online: 14 October 2014

Combinatorics, Probability and Computing (2015) 24, 521–527. © Cambridge University Press 2014
doi:10.1017/S0963548314000546

Maximizing the Number of Independent Sets of a Fixed Size

WENYING GAN[1], PO-SHEN LOH[2†] and BENNY SUDAKOV[1‡]

[1]Department of Mathematics, ETH, 8092 Zurich, Switzerland
(e-mail: ganw@math.ethz.ch, benjamin.sudakov@math.ethz.ch)

[2]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
(e–mail: ploh@cmu.edu)

Received 4 November 2013; revised 4 June 2014; first published online 14 October 2014

Let $i_t(G)$ be the number of independent sets of size t in a graph G. Engbers and Galvin asked how large $i_t(G)$ could be in graphs with minimum degree at least δ. They further conjectured that when $n \geqslant 2\delta$ and $t \geqslant 3$, $i_t(G)$ is maximized by the complete bipartite graph $K_{\delta,n-\delta}$. This conjecture has recently drawn the attention of many researchers. In this short note, we prove this conjecture.

2010 *Mathematics subject classification*: Primary 05C69
Secondary 05D99

1. Introduction

Given a finite graph G, let $i_t(G)$ be the number of independent sets of size t in a graph, and let $i(G) = \sum_{t \geqslant 0} i_t(G)$ be the total number of independent sets. There are many extremal results on $i(G)$ and $i_t(G)$ over families of graphs with various degree restrictions. Kahn [6] and Zhao [12] studied the maximum number of independent sets in a d-regular graph. Relaxing the regularity constraint to a minimum degree condition, Galvin [5] conjectured that the number of independent sets in an n-vertex graph with minimum degree $\delta \leqslant n/2$ is maximized by a complete bipartite graph $K_{\delta,n-\delta}$. This conjecture was recently proved (in stronger form) by Cutler and Radcliffe [3] for all n and δ, and they characterized the extremal graphs for $\delta > n/2$ as well.

One can further strengthen Galvin's conjecture by asking whether the extremal graphs also simultaneously maximize the number of independent sets of size t, for all t. This claim unfortunately is too strong, as there are easy counterexamples for $t = 2$. On the

† Research supported in part by NSF grant DMS-1201380 and by a USA–Israel BSF grant.
‡ Research supported in part by SNSF grant 200021-149111 and by a USA–Israel BSF grant.

other hand, no such examples are known for $t \geqslant 3$. Moreover, in this case Engbers and Galvin [4] made the following conjecture.

Conjecture 1.1. *For every $t \geqslant 3$ and $\delta \leqslant n/2$, the complete bipartite graph $K_{\delta,n-\delta}$ maximizes the number of independent sets of size t, over all n-vertex graphs with minimum degree at least δ.*

Engbers and Galvin [4] proved this for $\delta = 2$ and $\delta = 3$, and for all $\delta > 3$ they proved it when $t \geqslant 2\delta + 1$. Alexander, Cutler, and Mink [1] proved it for the entire range of t for bipartite graphs, but it appeared non-trivial to extend the result to general graphs. The first result for all graphs and all t was obtained by Law and McDiarmid [9], who proved the statement for $\delta \leqslant n^{1/3}/2$. This was improved by Alexander and Mink [2], who required that

$$\frac{(\delta+1)(\delta+2)}{3} \leqslant n.$$

In this short note, we completely resolve this conjecture.

Theorem 1.2. *Let $\delta \leqslant n/2$. For every $t \geqslant 3$, every n-vertex graph G with minimum degree at least δ satisfies $i_t(G) \leqslant i_t(K_{\delta,n-\delta})$, and when $t \leqslant \delta$, $K_{\delta,n-\delta}$ is the unique extremal graph.*

2. Proof

We will work with the complementary graph, and count cliques instead of independent sets. Cutler and Radcliffe [3] also discovered that the complement was more naturally amenable to extension; we will touch on this in our concluding remarks. Let us define some notation for use in our proof. A *t-clique* is a clique with t vertices. For a graph $G = (V, E)$, $\overline{G}$ is its complement, and $k_t(G)$ is the number of t-cliques in G. For any vertex $v \in V$, $N(v)$ is the set of the neighbours of v, $d(v)$ is the degree of v, and $k_t(v)$ is the number of t-cliques which contain vertex v. Note that $\sum_{v\in V} k_t(v) = tk_t(G)$. We also define $G + H$ as the graph consisting of the disjoint union of two graphs G and H. By considering the complementary graph, it is clear that our main theorem is equivalent to the following statement.

Proposition 2.1. *Let $1 \leqslant b \leqslant \Delta + 1$. For all $t \geqslant 3$, $k_t(G)$ is maximized by $K_{\Delta+1} + K_b$, over $(\Delta + 1 + b)$-vertex graphs with maximum degree at most Δ. When $t \leqslant b$, this is the unique extremal graph, and when $b < t \leqslant \Delta + 1$, the extremal graphs are $K_{\Delta+1} + H$, where H is an arbitrary b-vertex graph.* □

Remark. When $b \leqslant 0$, the number of t-cliques in graphs with maximum degree at most Δ is trivially maximized by the complete graph. On the other hand, when $b > (\Delta + 1)$, the problem becomes much more difficult, and our investigation is still ongoing. This paper focuses on the first complete segment $1 \leqslant b \leqslant \Delta + 1$, which, as mentioned in the Introduction, was previously attempted in [2, 4, 9].

Although our result holds for all $t \geqslant 3$, it turns out that the main step is to establish it for the case $t = 3$ using induction and double-counting. Afterwards, a separate argument will reduce the general $t > 3$ case to this case of $t = 3$.

Lemma 2.2. *Proposition 2.1 is true when* $t = 3$.

Proof. We proceed by induction on b. The base case $b = 0$ is trivial. Now assume it is true for $b - 1$. Suppose first that $k_3(v) \leqslant \binom{b-1}{2}$ for some vertex v. Applying the inductive hypothesis to $G - v$, we see that

$$k_3(G) \leqslant k_3(G - v) + k_3(v) \leqslant \binom{\Delta+1}{3} + \binom{b-1}{3} + \binom{b-1}{2} \leqslant \binom{\Delta+1}{3} + \binom{b}{3},$$

and equality holds if and only if $G - v$ is optimal and $k_3(v) = \binom{b-1}{2}$. By the inductive hypothesis, $G - v$ is $K_{\Delta+1} + H'$, where H' is a $(b-1)$-vertex graph. The maximum degree restriction forces v's neighbours to be entirely in H', and so $G = K_{\Delta+1} + H$ for some b-vertex graph H. Moreover, since $k_3(v) = \binom{b-1}{2}$ we get that for $b \geqslant 3$, H is a clique.

This leaves us with the case where $k_3(v) > \binom{b-1}{2}$ for every vertex v, which forces $b \leqslant d(v) \leqslant \Delta$. We will show that here, the number of 3-cliques is strictly suboptimal. The number of triples (u, v, w) where uv is an edge and vw is not an edge is clearly $\sum_{i=1}^{n} d(v)(n - 1 - d(v))$. Also, every set of three vertices either contributes 0 to this sum (if either all or none of the three edges between them are present), or contributes 2 (if they induce exactly one or exactly two edges). Therefore, we obtain an equality which also appeared in [11]:

$$2\left[\binom{n}{3} - (k_3(G) + k_3(\overline{G}))\right] = \sum_{v \in V} d(v)(n - 1 - d(v)).$$

Rearranging this equality and applying $k_3(\overline{G}) \geqslant 0$, we find

$$k_3(G) \leqslant \binom{n}{3} - \frac{1}{2}\sum_{v \in V} d(v)(n - 1 - d(v)). \tag{2.1}$$

Since we have already bounded $b \leqslant d(v) \leqslant \Delta$, and $b + \Delta = n - 1$ by definition, we have

$$d(v)(n - 1 - d(v)) \geqslant b\Delta.$$

Substituting this into (2.1) and using $n = (\Delta + 1) + b$,

$$k_3(G) \leqslant \binom{n}{3} - \frac{nb\Delta}{2} = \binom{\Delta+1}{3} + \binom{b}{3} - \frac{b(\Delta + 1 - b)}{2} < \binom{\Delta+1}{3} + \binom{b}{3},$$

because $b \leqslant \Delta$. This completes the case where every vertex has $k_3(v) > \binom{b-1}{2}$. $\square$

We reduce the general case to the case of $t = 3$ via the following variant of the celebrated Kruskal–Katona theorem [7, 8], which appears as Exercise 31b in Chapter 13 of Lovász's book [10]. Here, the generalized binomial coefficient $\binom{x}{k}$ is defined to be the product

$$\frac{1}{k!}(x)(x-1)(x-2)\cdots(x-k+1),$$

which exists for non-integral x.

Theorem 2.3. *Let $k \geqslant 3$ be an integer, and let $x \geqslant k$ be a real number. Then, every graph with exactly $\binom{x}{2}$ edges contains at most $\binom{x}{k}$ cliques of order k.*

We now use Lemma 2.2 and Theorem 2.3 to finish the general case of Proposition 2.1.

Lemma 2.4. *If Proposition 2.1 is true for $t = 3$, then it is also true for $t > 3$.*

Proof. Fix any $t \geqslant 4$. We proceed by induction on b. The base case $b = 0$ is trivial. For the inductive step, assume the result is true for $b - 1$. If there is a vertex v such that $k_3(v) \leqslant \binom{b-1}{2}$, then by applying Theorem 2.3 to the subgraph induced by $N(v)$, we find that there are at most $\binom{b-1}{t-1}$ cliques of order $t - 1$ entirely contained in $N(v)$. The t-cliques which contain v correspond bijectively to the $(t - 1)$-cliques in $N(v)$, and so $k_t(v) \leqslant \binom{b-1}{t-1}$. The same argument used at the beginning of Lemma 2.2 then correctly establishes the bound and characterizes the extremal graphs.

If some $k_3(v) = \binom{\Delta}{2}$, then the maximum degree condition implies that the graph contains a $K_{\Delta+1}$ which is disconnected from the remaining $b \leqslant \Delta + 1$ vertices, and the result also easily follows. Therefore, it remains to consider the case where all $\binom{b-1}{2} < k_3(v) < \binom{\Delta}{2}$, in which we will prove that the number of t-cliques is strictly suboptimal. It is well known and standard that for each fixed k, the binomial coefficient $\binom{x}{k}$ is strictly convex and increasing in the real variable x on the interval $x \geqslant k - 1$. Hence, $\binom{k}{k} = 1$ implies that $\binom{x}{k} < 1$ for all $k - 1 < x < k$, and so Theorem 2.3 then actually applies for all $x \geqslant k - 1$. Thus, if we define $u(x)$ to be the positive root of $\binom{u}{2} = x$, that is,

$$u(x) = \frac{(1 + \sqrt{1 + 8x})}{2},$$

and let

$$f_t(x) = \begin{cases} 0 & \text{if } u(x) < t - 2, \\ \binom{u(x)}{t-1} & \text{if } u(x) \geqslant t - 2, \end{cases} \tag{2.2}$$

the application of the Kruskal–Katona theorem in the previous paragraph establishes that $k_t(v) \leqslant f_t(k_3(v))$.

We will also need that $f_t(x)$ is strictly convex for $x > \binom{t-2}{2}$. For this, observe that by the generalized product rule,

$$f_t'(x) = u' \cdot [(u - 1)(u - 2) \cdots (u - (t - 2)) + \cdots + u(u - 1) \cdots (u - (t - 3))],$$

which is $u'(x)$ multiplied by a sum of $t - 1$ products. Since

$$u'(x) = \frac{2}{\sqrt{1 + 8x}},$$

for any constant C,

$$(u')(u - C) = 1 - \frac{(2C - 1)}{\sqrt{1 + 8x}}.$$

Note that this is a positive increasing function when $C \in \{1,2\}$ and $x > \binom{t-2}{2}$. In particular, since $t \geqslant 4$, each of the $t-1$ products contains a factor of $(u-1)$ or $(u-2)$, or possibly both; we can then always select one of them to absorb the (u') factor, and conclude that $f_t'(x)$ is the sum of $t-1$ products, each of which is composed of $t-2$ factors that are positive increasing functions on $x > \binom{t-2}{2}$. Thus $f_t(x)$ is strictly convex on that domain, and since $f_t(x) = 0$ for $x \leqslant \binom{t-2}{2}$, it is convex everywhere.

If $t = \Delta + 1$, there will be no t-cliques in G unless G contains a $K_{\Delta+1}$, which must be isolated because of the maximum degree condition; we are then finished as before. Hence we may assume $t \leqslant \Delta$ for the remainder, which in particular implies that $f_t(x)$ is strictly convex and strictly increasing in the neighbourhood of $x = \binom{\Delta}{2}$. Let the vertices be $v_1, \ldots, v_n$, and define $x_i = k_3(v_i)$. We have

$$tk_t(G) = \sum_{v \in V} k_t(v) \leqslant \sum_{i=1}^{n} f_t(x_i),$$

and so it suffices to show that

$$\sum f_t(x_i) < t\binom{\Delta+1}{t} + t\binom{b}{t}$$

under the following conditions, the latter of which comes from Lemma 2.2:

$$\binom{b-1}{2} < x_i < \binom{\Delta}{2}, \quad \sum_{i=1}^{n} x_i \leqslant 3\binom{\Delta+1}{3} + 3\binom{b}{3}. \tag{2.3}$$

To this end, consider a tuple of real numbers $(x_1, \ldots, x_n)$ which satisfies the conditions. Although (2.3) constrains each x_i within an open interval, we will perturb the x_i within the closed interval which includes the endpoints, in such a way that the objective $\sum f_t(x_i)$ is non-decreasing, and we will reach a tuple which achieves an objective value of exactly $t\binom{\Delta+1}{t} + t\binom{b}{t}$. Finally, we will use our observation of strict convexity and monotonicity around $x = \binom{\Delta}{2}$ to show that one of the steps strictly increased $\sum f_t(x_i)$, which will complete the proof.

First, since the upper limit for $\sum x_i$ in (2.3) is achievable by setting $\Delta + 1$ of the x_i to $\binom{\Delta}{2}$ and b of the x_i to $\binom{b-1}{2}$, and $f_t(x)$ is non-decreasing, we may increase some of the x_i in the interval

$$\binom{b-1}{2} \leqslant x_i \leqslant \binom{\Delta}{2}$$

such that the inequality for $\sum x_i$ in (2.3) is tight. Next, by convexity of $f_t(x)$, we may push apart x_i and x_j while conserving their sum, and the objective is non-decreasing. After a finite number of steps, we arrive at a tuple in which all but at most one of the x_i is equal to either the lower limit $\binom{b-1}{2}$ or the upper limit $\binom{\Delta}{2}$, and

$$\sum x_i = 3\binom{\Delta+1}{3} + 3\binom{b}{3}.$$

However, since this value of $\sum x_i$ is achievable by $\Delta + 1$ many $\binom{\Delta}{2}$ and b many $\binom{b-1}{2}$, this implies that in fact, the tuple of the x_i has precisely this form. (To see this, note that by an affine transformation, the statement is equivalent to the fact that if n and k are integers, and $0 \leqslant y_i \leqslant 1$ are n real numbers which sum to k, all but one of which is at an endpoint, then exactly k of the y_i are equal to 1 and the rest are equal to 0.) Thus, our final objective is equal to

$$(\Delta + 1)\binom{\Delta}{t-1} + b\binom{b-1}{t-1} = t\binom{\Delta+1}{t} + t\binom{b}{t},$$

as claimed. Finally, since some x_i take the value $\binom{\Delta}{2}$, the strictness of $f_t(x)$'s monotonicity and convexity in the neighbourhood $x = \binom{\Delta}{2}$ implies that at some stage of our process, we strictly increased the objective. Therefore, in this case where all

$$\binom{b-1}{2} < k_3(v) < \binom{\Delta}{2},$$

the number of t-cliques is indeed sub-optimal, and our proof is complete. □

3. Concluding remarks

The natural generalization of Proposition 2.1 considers the maximum number of t-cliques in graphs with maximum degree Δ and $n = a(\Delta + 1) + b$ vertices, where $0 \leqslant b < \Delta + 1$. In the language of independent sets, this question was also proposed by Engbers and Galvin [4]. The case $a = 0$ is trivial, and Proposition 2.1 completely solves the case $a = 1$. We believe that also for $a > 1$ and $t \geqslant 3$, $k_t(G)$ is maximized by $aK_{\Delta+1} + K_b$, over $(a(\Delta + 1) + b)$-vertex graphs with maximum degree at most Δ.

An easy double-counting argument shows that it is true when $b = 0$. When $a \geqslant 2$ and $b > 0$, the problem seems considerably more delicate. Nevertheless, the same proof that we used in Lemma 2.4 (*mutatis mutandis*) shows that the general case $t > 3$ of this problem can be reduced to the case $t = 3$. Therefore, the most intriguing and challenging part is to show that $aK_{\Delta+1} + K_b$ maximizes the number of triangles over all graphs with $(a(\Delta + 1) + b)$ vertices and maximum degree at most Δ. We have some partial results on this main case, but our investigation is still ongoing.

References

[1] Alexander, J., Cutler, J. and Mink, T. (2012) Independent sets in graphs with given minimum degree. *Electron. J. Combin.* **19** P37.

[2] Alexander, J. and Mink, T. (2013) A new method for enumerating independent sets of a fixed size in general graphs. `arXiv:1308.3242`

[3] Cutler, J. and Radcliffe, A. J. (2013) The maximum number of complete subgraphs in a graph with given maximum degree. *J. Combin. Theory Ser. B* **104** 60–71.

[4] Engbers, J. and Galvin, D. (2014) Counting independent sets of a fixed size in graphs with a given minimum degree. *J. Graph Theory* **76** 149–168.

[5] Galvin, D. (2011) Two problems on independent sets in graphs. *Discrete Math.* **311** 2105–2112.

[6] Kahn, J. (2001) An entropy approach to the hard-core model on bipartite graphs. *Combin. Probab. Comput.* **10** 219–237.

[7] Katona, G. (1968) A theorem of finite sets. In *Theory of Graphs: Proc. Colloq., Tihhany, 1966*, Academic Press, pp. 187–207.

[8] Kruskal, J. B. (1964) The number of simplices in a complex. In *Mathematical Optimization Techniques* (R. Bellman, ed.), University of California Press, pp. 251–278.
[9] Law, H. and McDiarmid, C. (2013) On independent sets in graphs with given minimum degree. *Combin. Probab. Comput.* **22** 874–884.
[10] Lovász, L. (2007) *Combinatorial Problems and Exercises*, AMS.
[11] Sauvé, I. L. (1961) On chromatic graphs. *Amer. Math. Monthly* **68** 107–111.
[12] Zhao, Y. (2010) The number of independent sets in a regular graph. *Combin. Probab. Comput.* **19** 109–112.