Large induced subgraphs with k vertices of almost maximum degree

António Girão * Kamil Popielarz †

June 23, 2021

Abstract

In this note we prove that for every integer k, there exist constants $g_1(k)$ and $g_2(k)$ such that the following holds. If G is a graph on n vertices with maximum degree Δ then it contains an induced subgraph H on at least $n - g_1(k)\sqrt{\Delta}$ vertices, such that H has k vertices of the same degree of order at least $\Delta(H) - g_2(k)$. This solves a conjecture of Caro and Yuster up to the constant $g_2(k)$.

1 Introduction

Given a graph G, let the repetition number, denoted by $rep(G)$, be the maximum multiplicity of a vertex degree. Trivially, any graph G of order at least two contains at least two vertices of the same degree, i.e. $rep(G) \geq 2$. This parameter has been widely studied by several researchers (e.g., [2, 4, 7, 9, 10]), in particular, by Bollobás and Scott, who showed that for every $k \geq 2$ there exist triangle-free graphs on n vertices with $rep(G) \leq k$ for which $\alpha(G) = (1 + o(1))n/k$ ([4]). As there are infinitely many graphs having repetition number two, it is natural to ask what is the smallest number of vertices one needs to delete from a graph in order to increase the repetition number of the remaining induced subgraph. This question was partially answered by Caro, Shapira and Yuster in [6], indeed, they proved that for every k there exists a constant $C(k)$ such that given any graph on n vertices one needs to remove at most $C(k)$ vertices and thus obtain an induced subgraph with at least $\min\{k, n-C(k)\}$ vertices of the same degree. Related to this question, Caro and Yuster ([8]) considered the problem of finding the largest induced subgraph H of a graph G which contains at least k vertices of degree $\Delta(H)$. To do so they defined $f_k(G)$ to be the smallest number of vertices one needs to remove from a graph G such that the remaining induced subgraph has its maximum degree attained by at least k vertices. They found examples of graphs on n vertices for which $f_2(G) \geq (1 - o(1))\sqrt{n}$ and conjectured $f_k(G) \leq O(\sqrt{n})$ for every graph G on n vertices. In the same paper they established the conjecture for $k \leq 3$.

*Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, UK; A.Girao@dpmms.cam.ac.uk

†Department of Mathematics, University of Memphis, Memphis, Tennessee; kamil.popielarz@gmail.com

The following more general conjecture was posed recently by Caro, Lauri and Zarb in [5].

Conjecture 1. *For every $k \geq 2$ there is a constant $g(k)$ such that given a graph G with maximum degree Δ, one can remove at most $g(k)\sqrt{\Delta}$ vertices such that the remaining subgraph $H \subseteq G$ has at least k vertices of degree $\Delta(H)$.*

Let us define $g(k, \Delta) = \max\{f_k(G) : \Delta(G) \leq \Delta\}$. In the same paper, they proved that $g(2, \Delta) = \lceil \frac{3+\sqrt{8\Delta+1}}{2} \rceil$ and stated that $g(3, \Delta) \leq 42\sqrt{\Delta}$. We should point out that, if true, the conjecture is best possible, as there are graphs on n vertices found in [5] for which any induced subgraph on more than $n - \frac{k}{2}\sqrt{\Delta}$ does not contain k vertices of the same maximum degree. We shall present such constructions in Section 3.

In this note we prove the following approximate version of Conjecture 1.

Theorem 1. *For every positive integer k, there exist constants $g_1(k)$ and $g_2(k)$ such that the following holds. If G is a graph on n vertices with maximum degree Δ then it contains an induced subgraph H on at least $n - g_1(k)\sqrt{\Delta}$ vertices, such that H has k vertices of the same degree at least $\Delta(H) - g_2(k)$.*

2 Proofs

Given a partition of $\{1, 2, \ldots, n\}$ into t sets, $A_1, A_2 \ldots, A_t$, and a strictly decreasing sequence of non-negative integers $r_1 > r_2 > r_3 > \ldots > r_t$, we say $\mathcal{A}$ is an $(r_1, r_2, \ldots, r_t)$-uniform cover of $\{1, 2 \ldots, n\}$ if $\mathcal{A}$ is a multiset of subsets of $\{1, 2, \ldots n\}$ such that, whenever $i \in \{1, \ldots, t\}$ and $j \in A_i$, we have $|\{A \in \mathcal{A} : j \in A\}| = r_i$. Note that $\mathcal{A}$ is a multiset, hence we allow repetitions.

We call an $(r_1, r_2, \ldots, r_t)$-uniform cover $\mathcal{A}$ of $\{1, 2, \ldots, n\} = A_1 \cup A_2 \cup \ldots \cup A_t$ *irreducible* if there is no proper $(r'_1, \ldots, r'_t)$-uniform cover $\mathcal{B} \subset \mathcal{A}$, for some strictly decreasing sequence of non-negative integers $r'_1 > r'_2 > \ldots > r'_t$.

Given a uniform cover $\mathcal{A}$ of $\{1, 2, \ldots, n\}$ and a subset $B \subseteq \{1, 2, \ldots, n\}$ we define $w_{\mathcal{A}}(B)$ to be the number of times B appears in $\mathcal{A}$.

Lemma 2. *For all $n \in \mathbb{N}$, there exists $f(n)$ such that for any $1 \leq t \leq n$ and any partition of $\{1, 2, \ldots, n\}$ into t sets $A_1, A_2, \ldots, A_t$, every $(r_1, r_2, \ldots, r_t)$-uniform cover $\mathcal{A}$ of $\{1, 2, \ldots, n\}$ contains a proper $(r'_1, r'_2, \ldots, r'_t)$-uniform sub-cover $\mathcal{B} \subset \mathcal{A}$ with $r'_1 \leq f(n)$.*

Proof. We shall prove there are only finitely many *irreducible* covers. For otherwise, let us assume there exists an infinite sequence $\{B_i\}_{i \in \mathbb{N}}$ of *irreducible* uniform covers. Since there are only finitely many partitions of $\{1, 2, \ldots, n\}$, we may pass to an infinite subsequence $\{B_{l_i}\}_{i \in \mathbb{N}}$ of uniform covers of the same partition of $\{1, 2, \ldots, n\}$. Now, choose $A \subseteq \{1, 2, \ldots, n\}$ and consider the sequence of non-negative integers $\{w_{B_{l_i}}(A)\}_{i \in \mathbb{N}}$, clearly it must contain an infinite non-decreasing subsequence $w_{B_{l_{i_1}}}(A) \leq w_{B_{l_{i_2}}}(A) \leq \ldots$. We restrict our attention to this subsequence of the uniform covers $B_{l_{i_1}}, B_{l_{i_2}}, \ldots$ and iteratively apply the same argument for the remaining subsets of $\{1, 2, \ldots, n\}$, always passing to a subsequence of the previous sequence of uniform covers. After we have done it for every subset of $\{1, 2, \ldots, n\}$, we must end up with two distinct *irreducible* uniform covers (actually an infinite sequence) $\mathcal{A}, \mathcal{B}$ for which $w_{\mathcal{A}}(F) \leq w_{\mathcal{B}}(F)$ for every $F \subseteq \{1, 2, \ldots, n\}$.

This implies $\mathcal{A} \subseteq \mathcal{B}$, which is a contradiction. Take $f(n)$ to be the maximum r_1 over all *irreducible* uniform covers of $\{1, 2, \ldots, n\}$. $\square$

Lemma 3. *For every $n \in \mathbb{N}$, there exists $f(n)$ such that the following holds. Let $G = (A, B)$ be a bipartite graph with $A = \{x_1, x_2, \ldots, x_n\}$. Then there exists a subset $W \subseteq V(B)$ of size at most $n \cdot f(n) = f'(n)$, such that the induced bipartite graph $G' = G[A, (B \setminus W)]$ has the property that*

$$\textit{if } d_G(x_i) > d_G(x_j) \textit{ then } d_G(x_i) - d_{G'}(x_i) > d_G(x_j) - d_{G'}(x_j).$$

Proof. Partition A into $A_1, \ldots, A_t$, so that two vertices belong to the same part if they have the same degree. Let r_i be the degree of the vertices in A_i. We may assume that $r_1 > r_2 > \cdots > r_t$. The lemma follows as a corollary of Lemma 2. Indeed, for every vertex $w \in B$, let $A_w \subseteq \{1, 2, \ldots, n\}$ such that $i \in A_w$ if x_i is a neighbour of w in G. Note that $\mathcal{A} = \{A_w : w \in B\}$ is an $(r_1, r_2, \ldots, r_t)$-uniform cover of $\{1, 2, \ldots, n\}$. Applying now Lemma 2, we can find a $(r'_1, r'_2, \ldots, r'_t)$-uniform sub-cover $\mathcal{B} \subseteq \mathcal{A}$ with $r'_1 \leq f(n)$. Let $W = \{w \in B : A_w \in \mathcal{B}\}$ and $G' = G[A, (B \setminus W)]$. It is easy to see that $|W| \leq n \cdot f(n)$ and that the property is satisfied by the definition of uniform cover. $\square$

Given a positive integer k and a graph G with the vertex set $\{x_1, \ldots, x_n\}$ such that $d(x_1) \geq \cdots \geq d(x_n)$, let $r_k(G) := \Delta(G) - d_G(x_k)$ be the difference between the maximum degree and the degree of vertex x_k.

Theorem 4. *For every positive integer k there exists $h(k)$ such that the following holds. If G is a graph on n vertices with maximum degree Δ then it contains an induced subgraph H on at least $n - (h(k) + k)\sqrt{\Delta}$ vertices, such that $r_k(H) \leq h(k) \cdot k$.*

Proof. The proof consists of two parts. Firstly, we shall show that we can remove at most $k\sqrt{\Delta}$ vertices from G so that in the remaining graph H' we have $r_k(H') \leq \sqrt{\Delta}$. Then we iteratively apply Lemma 3 (at most $\sqrt{\Delta}$ times) in order to obtain an induced subgraph H of H' on at least $n - (h(k) + k)\sqrt{\Delta}$ vertices such that $r_k(H) \leq h(k) \cdot k$. We may take $h(k)$ to be $f'(k)$ from Lemma 3.

We start by showing there is a large induced subgraph $H' \subseteq G$ with $r_k(H') \leq \sqrt{\Delta}$.

Claim 1. *There is an induced subgraph H' of G on at least $n - k\sqrt{\Delta}$ vertices such that $r_k(H') \leq \sqrt{\Delta}$.*

Proof of Claim 1. Consider the following procedure. Let $G_0 = G$ and suppose that $G_0 \supset \cdots \supset G_i$ have been defined. If G_i does not have the required property then let G_{i+1} be obtained from G_i by removing k vertices with largest degrees in G_i. Notice that $\Delta(G_{i+1}) \leq \Delta(G_i) - \sqrt{\Delta}$ and $|G_{i+1}| = |G_i| - k$. Observe that the procedure will stop after at most $\sqrt{\Delta}$ steps, as otherwise the obtained graph would have maximum degree 0. Since $|G_i| \geq n - i \cdot k$ we have that $|H'| \geq n - k\sqrt{\Delta}$. $\square$

We now proceed to the second part of the proof and iteratively apply Lemma 3. In each step we remove at most $h(k)$ vertices from H' while decreasing the value of r_k and we stop when r_k is at most $k \cdot h(k)$.

Let $H_0 = H'$ and suppose that $H_0, \ldots, H_i$ have already been defined. If $r_k(H_i) \leq k \cdot h(k)$ then we are done, so we may assume that $r_k(H_i) > k \cdot h(k)$. Let $A = \{x_1, \ldots, x_k\}$ be a set of k vertices with the largest degrees in H_i and write B for $H_i \setminus A$. Without loss of generality we may assume that $d_{H_i}(x_1) \geq \cdots \geq d_{H_i}(x_k)$. Since $r_k(H_i) \geq k \cdot h(k)$ there must exist $l \in \{2, \ldots, k\}$ such that $d_{H_i}(x_l) > d_{H_i}(x_{l-1}) + h(k)$. Now consider the bipartite subgraph $K = H_i[A, B]$. By Lemma 3, with $G = K$ and $n = k$, we can remove a set $W \subset B$ of at most $f'(k) = h(k)$ vertices from B, and obtain $K' = H_i[A, (B \setminus W)]$ such that

$$\text{for any } x, y \in A, \text{ if } d_K(x) < d_K(y) \text{ then } d_K(x) - d_{K'}(x) < d_K(y) - d_{K'}(y). \tag{1}$$

Let $H_{i+1} = H_i \setminus W$ (hence $|H_{i+1}| \geq |H_i| - |W| \geq |H_i| - h(k)$). The following claim asserts that the above procedure will stop after at most $\sqrt{\Delta}$ steps.

Claim 2. *$r_k(H_{i+1}) < r_k(H_i)$.*

Proof of Claim 2. Let z be a vertex with the maximum degree and w a vertex with the k'th largest degree in H_{i+1}. Observe that $z = x_t$ for some $t \geq l$ and $d_{H_{i+1}}(w) \geq d_{H_{i+1}}(x_s)$ for some $s < l$. First, notice that $d_{H_i}(x_t) - d_{H_i}(x_s) \leq d_{H_i}(x_1) - d_{H_i}(x_k) = r_k(H_i)$. Hence, $r_k(H_{i+1}) = d_{H_{i+1}}(z) - d_{H_{i+1}}(w) \leq d_{H_{i+1}}(x_t) - d_{H_{i+1}}(x_s) < d_{H_i}(x_t) - d_{H_i}(x_s) \leq r_k(H_i)$, where the strict inequality follows from (1) since $d_K(x_t) > d_K(x_s)$. □

As in each iteration the value of r_k decreases, we must stop after at most $r_k(H') = \sqrt{\Delta}$ steps thus getting an induced subgraph $H \subset H'$ with $r_k(H) \leq k \cdot h(k)$ and $|H| \geq |H'| - h(k)\sqrt{\Delta} \geq n - (h(k) + k)\sqrt{\Delta}$. □

In order to prove Theorem 1 we need the following theorem of Caro, Shapira and Yuster, appearing in [6], whose proof is inspired by the one used by Alon and Berman in [1].

Theorem 5. *For positive integers r, d, q, the following holds. Any sequence of $n \geq (\lceil q/r \rceil + 2)(2rd + 1)^d$ elements of $[-r, r]^d$ whose sum, denoted by z, is in $[-q, q]^d$ contains a subsequence of length at most $(\lceil q/r \rceil + 2)(2rd + 1)^d$ whose sum is z.*

As usual, we write $R(k)$ (see e.g. [3]) for the *two coloured Ramsey number*, the least integer n such that in any two colouring of the edges of the complete graph on n vertices, there is a monochromatic K_k.

Proof of Theorem 1. Firstly, we apply Theorem 4 with $k = R(k)$ to find a large induced subgraph $G' \subset G$ of order at least $n' \geq n - (h(R(k)) + R(k))\sqrt{\Delta}$ and with vertex set $\{x_1, \ldots, x_{n'}\}$ where $d(x_1) \geq d(x_2) \geq \cdots \geq d(x_{n'})$ and $d(x_1) - d(x_{R(k)}) \leq h(R(k)) \cdot R(k) = M$. Now we follow the proof of Theorem 1.1 in [6].

By the definition of $R(k)$ we can find a set S of k vertices in $\{x_1, \ldots, x_{R(k)}\}$ that induces either a complete graph or an independent set.

Without loss of generality, assume that $S = \{v_{n'-k+1}, \ldots, v_{n'}\}$ and $V(G) \setminus S = \{v_1, \ldots, v_{n'-k}\}$. Let $e(v_i, v_j)$ be equal to 1 if there is an edge between v_i and v_j, and 0 otherwise. We construct a sequence X of $n' - k$ vectors $w_1, \ldots, w_{n'-k}$ in $[-1, 1]^{k-1}$ as follows. The coordinate j of w_i is $e(v_{n'-k+j}, v_i) - e(v_{n'}, v_i)$ for $i = 1, \ldots, n' - k$ and

$j = 1, \ldots, k-1$. It is clear that $e(v_{n'-k+j}, v_i) - e(v_{n'}, v_i) \in [-1, 1]$ as required. Consider the sum of all the j'th coordinates,

$$\begin{aligned}\sum_{i=1}^{n'-k} \left(e(v_{n'-k+j}, v_i) - e(v_{n'}, v_i)\right) &= \sum_{i=1}^{n'-k} e(v_{n'-k+j}, v_i) - \sum_{i=1}^{n'-k} e(v_{n'}, v_i) \\ &= (d(v_{n'-k+j}) - a) - (d(v_{n'}) - a) = d(v_{n'-k+j}) - d(v_{n'}) \\ &\leq M,\end{aligned}$$

where $a = k-1$ if $G'[S]$ is complete, and $a = 0$ otherwise. Hence,

$$z = \sum_{i=1}^{n'-k} w_i \in [-M, M]^{k-1} .$$

By Theorem 5, with $d = k-1$ and $q = M$, there is a subsequence of X of size at most $(M+2)(2k-1)^{k-1}$ whose sum is z. Deleting the vertices of G' corresponding to the elements of this subsequence results in an induced subgraph $H \subset G'$ in which all the k vertices of S have the same degree of order at least $\Delta(H) - \left(M + (M+2)(2k-1)^{k-1}\right)$. Choosing $g_1(k) = g_2(k) = h(R(k))(4k)^k$ we conclude the theorem. □

3 Remarks

In the previous section, we proved that every graph contains a large induced subgraph with at least k vertices having the same degree of order almost the maximum degree. Note that Theorem 1 is sharp up to the size of the functions $g_1(k)$ and $g_2(k)$. Indeed, there are graphs for which one needs to remove "roughly" $\frac{k}{2}\sqrt{\Delta}$ vertices to force the remaining subgraph to have k vertices with the same degree "near" the maximum degree. For any k and Δ, let G^{Δ} be the disjoint union of the stars $K_{1,n_1}, \ldots, K_{1,n_t}$, where $n_i = i \cdot \sqrt{\Delta}$, for $i \in \left\{1, \ldots, t = \sqrt{\Delta}\right\}$ and let G_k^{Δ} be the disjoint union of $k/2$ copies of G^{Δ}. It is easy to see that, for any constant D, one needs to remove at least $\frac{k}{2}\sqrt{\Delta} - \frac{k}{2}D$ vertices from G_k^{Δ} in order to obtain an induced graph H with k vertices of the same degree of order at least $\Delta(H) - D$.

Whether removing $C(k)\sqrt{\Delta}$ vertices is enough to force the remaining induced subgraph to have at least k vertices attaining the maximum degree remains an interesting open question.

4 Acknowledgments

This project was carried through during our stay in IMT School for Advanced Studies Lucca. We would like to thank the Institute for their kind hospitality.

References

[1] Alon, N., and Berman, K. Regular hypergraphs, Gordon's lemma, Steinitz' lemma and invariant theory. *Journal of Combinatorial Theory, Series A 43* (1986), 91–97.

[2] Balister, P., Bollobás, B., Lehel, J., and Morayne, M. Repeated degrees in random uniform hypergraphs. *SIAM Journal on Discrete Mathematics 27*, 1 (2013), 145–154.

[3] Bollobás, B. *Modern Graph Theory*. Springer, 1998.

[4] Bollobás, B., and Scott, A. Independent sets and repeated degrees. *Discrete Mathematics 170*, 1 (1997), 41–49.

[5] Caro, Y., Lauri, Y., and Zarb, C. Equating two maximum degrees. *ArXiv e-prints* (apr 2017).

[6] Caro, Y., Shapira, A., and Yuster, R. Forcing k-repetitions in degree sequences. *Electronic Journal of Combinatorics 16:#R7* (2013).

[7] Caro, Y., and West, D. Repetition number of graphs. *Electronic Journal of Combinatorics 16:#R7* (2009).

[8] Caro, Y., and Yuster, R. Large induced subgraphs with equated maximum degree. *Discrete Mathematics 310*, 4 (2010), 742–747.

[9] Chen, G., and Schelp, R. Ramsey problems with bounded degree spread. *Combinatorics, Probabilty and Computing 2* (1993), 263 – 269.

[10] Erdős, P., Chen, G., Rousseau, C., and Schelp, R. Ramsey problems involving degrees in edge coloured complete graphs of vertices belonging to monochromatic subgraphs. *European Journal of Combinatorics 14*, 3 (1993), 183 – 189.