A result on large induced subgraphs with prescribed residues in bipartite graphs

Zach Hunter

Mathematical Institute
Oxford University, Exeter College
Oxford, UK.

zachary.hunter@exeter.ox.ac.uk

Submitted: Aug 15, 2022; Accepted: Nov 7, 2022; Published: Jan 27, 2023

Abstract

It was proved by Scott that for every $k \geqslant 2$, there exists a constant $c(k) > 0$ such that for every bipartite n-vertex graph G without isolated vertices, there exists an induced subgraph H of order at least $c(k)n$ such that $\deg_H(v) \equiv 1 \pmod{k}$ for each $v \in H$. Scott conjectured that $c(k) = \Omega(1/k)$, which would be tight up to the multiplicative constant. We confirm this conjecture.

Mathematics Subject Classifications: 05C07, 05C35

1 Introduction

Given a graph G and integers $q > r \geqslant 0$, we define $f(G, r, q)$ to be the maximum order of an induced subgraph H of G where $\deg_H(v) \equiv r \pmod{q}$ for all $v \in H$ (or if no such H exists, we set $f(G, r, q) = 0$).

There are many questions and conjectures concerning the behavior of $f(G, r, q)$ for various G, r, q. An old unpublished result of Gallai in this area is that[1] $f(G, 0, 2) \geqslant n/2$ for every n-vertex graph (see [7, Excercise 5.17] for a proof). Further questions about the behavior of f received attention around 20-30 years ago (see e.g., [2, 3, 9, 10]). And more recently, this topic has had a renaissance (see e.g., [1, 5, 6, 8]).

This note will focus on an old result of Scott. For positive integer k, we define $c(k)$ to be $\inf_G\{f(G, 1, k)/|G|\}$ where G ranges over all bipartite graphs with $\delta(G) \geqslant 1$. The following was proved by Scott:

[1]Actually what Gallai proved was slightly stronger. He showed that for each graph G, we can partition $V(G)$ into two parts A, B so that $\deg_{G[A]}(v) \equiv 0 \pmod{2}$ (respectively $\deg_{G[B]}(v) \equiv 0 \pmod{2}$) for each $v \in A$ (respectively $v \in B$).

Theorem 1 ([10, Lemma 8]). *Let $k \geqslant 2$. Then*

$$1/(2^k + k + 1) \leqslant c(k) \leqslant 1/k.$$

Scott observed that a slightly more careful argument could further show that $c(k) = \Omega\left(\frac{1}{k^2 \log k}\right)$.

In this note we give an improved lower bound to $c(k)$ which is optimal up to the (implied) multiplicative constant.

Theorem 2. *Let $k \geqslant 2$. Then $c(k) = \Omega(1/k)$.*

This is done by taking the improved argument suggested by Scott, and then applying a dyadic pigeonhole argument which was previously overlooked.

2 Proof of Theorem 2

We will need the following result on the mixing time of random walks modulo k.

Lemma 3. *Let X_i be i.i.d. random variables that sample $\{0, 1\}$ uniformly at random. If $N \geqslant k^3$, then*[2] $\mathbb{P}\left(\sum_{i=1}^{N} X_i \equiv 1 \pmod{k}\right) \geqslant (1 - o_k(1))/k$.

Lemma 3 is a mild corollary of several known results, and we note k^3 could be replaced with $k^2 \log k$ (or any function which is $\omega(k^2)$).

For convenience, a fully elementary proof of Lemma 3 is provided in Appendix A.

In [10], when Scott outlined how to prove $c(k) \geqslant \Omega\left(\frac{1}{k^2 \log k}\right)$, he noted that Lemma 3 (the key to the improvement) can be derived by slightly modifying the argument in [4, Theorem 2 of Chapter 3]. These appropriate modifications now appear in [5]. Namely, the interested reader can confirm that Lemma 3 follows from the proof[3] of [5, Lemma 2.3]. Both of these proofs rely on discrete Fourier Analysis.

We now proceed to the main proof.

Proof of Theorem 2. Let G be an n-vertex bipartite graph with $\delta(G) \geqslant 1$, and let V_1, V_2 bipartition G with $|V_1| \geqslant |V_2|$. We shall write c_1, c_2 to denote small positive quantities which will be determined later (it would suffice to take $c_1 = 1/4, c_2 = 1/2$, but for clarity and a slightly better constant we will only consider their values at the end of the proof and shall have them depend slightly on k). Our proof splits into three cases.

We take $W_1 \subset V_2$ to be a minimal set satisfying $|N(v) \cap W_1| > 0$ for all $v \in V_1$ (i.e., W_1 is a minimal dominating set of V_1). By minimality of W_1, for each $w \in W_1$ there must exist $v_w \in V_1$ where $N(v_w) \cap W_1 = \{w\}$. Let $S_1 = \{v_w : w \in W_1\}$. We conclude that $W_1 \cup S_1$ induces a matching in G, proving that $f(G, 1, k) \geqslant 2|W_1|$.

Hence, we will be done if $|W_1| \geqslant c_1|V_1|/k$ (this is "Case 1"). So we continue assuming $|W_1| < c_1|V_1|/k$.

[2] Throughout this paper, we write $o_k(1)$ to denote quantities that tend to zero as $k \to \infty$.

[3] In [5], the statement of their lemma hides some constants which are necessary to verify our statement of Lemma 3.

For $2 \leqslant i \leqslant k-1$, we inductively create sets W_i, S_i. We take $W_i \subset W_{i-1}$ to be a minimal dominating set of $V_1 \setminus \left(\bigcup_{j=1}^{i-1} S_j\right)$. And like in the above, we take $S_i \subset V_1 \setminus \left(\bigcup_{j=1}^{i-1} S_j\right)$ so that $W_i \cup S_i$ induces a matching in G.

Let $T = V_1 \setminus \left(\bigcup_{i=1}^{k-1} S_i\right)$. We have

$$\begin{aligned} |T| &= |V_1| - \sum_{i=1}^{k-1} |S_i| \\ &= |V_1| - \sum_{i=1}^{k-1} |W_i| \\ &\geqslant |V_1| - (k-1)|W_1| \\ &\geqslant (1-c_1)|V_1|. \end{aligned}$$

Next, let $T^* = \{v \in T : |N(v) \cap W_{k-1}| \geqslant k^3\}$. Supposing that $|T^*| \geqslant c_2|V_1|$ (this is "Case 2"), we will deduce that $f(G,1,k) \geqslant (c_2 - o_k(1))|V_1|/k$.

Indeed, let $U \subset W_{k-1}$ be a random subset where each element is included (independently) with probability 1/2. We set $T_U = \{v \in T : |N(v) \cap U| \equiv 1 \pmod{k}\}$. By Lemma 3, we have that $\mathbb{P}(v \in T_U) \geqslant (1 - o_k(1))/k$ for each $v \in T^*$. Thus by linearity of expectation we may fix some $U \subset W_{k-1}$ where $|T_U| \geqslant |T^*|(1-o_k(1))/k \geqslant (c_2 - o_k(1))|V_1|/k$. Next choosing $S \subset \bigcup_{i=1}^{k-1} S_i$ so that $|N(u) \cap (T_U \cup S)| \equiv 1 \pmod{k}$ for each $u \in U$, we have that $S \cup U \cup T_U$ induces a subgraph in G demonstrating that $f(G,1,k) \geqslant |S \cup U \cup T_U| \geqslant |T_U| \geqslant (c_2 - o_k(1))|V_1|/k$.

Otherwise, we must have that $T \setminus T^*$, the set of $v \in T$ where $|N(v) \cap W_{k-1}| < k^3$, has $> (1 - c_1 - c_2)|V_1|$ elements (this is "Case 3"). By dyadic pigeonhole, there exists some $0 \leqslant p \leqslant \log(k^3) = O(\log k)$ so that

$$\begin{aligned} |\{v \in T : 2^p \leqslant |N(v) \cap W_{k-1}| < 2^{p+1}\}| &\geqslant |T \setminus T^*|/O(\log k) \\ &\geqslant (1 - c_1 - c_2)|V_1|/O(\log k). \end{aligned}$$

Take $T' = \{v \in T : 2^p \leqslant |N(v) \cap W_{k-1}| < 2^{p+1}\}$ to be this large set.

We let $U \subset W_{k-1}$ be a random subset so that each element is included (independently) with probability $1/2^p$. Defining T_U as before, some casework[4] shows $\mathbb{P}(v \in T_U) \geqslant e^{-2}$ for each $v \in T'$. Hence, by linearity of expectation, we may fix U so that $|T_U| \geqslant e^{-2}|T'|$. As above we may find $S \subset \bigcup_{i=1}^{k-1} S_i$ so that $S \cup U \cup T_U$ demonstrates that $f(G,1,k) \geqslant |S \cup U \cup T_U| \geqslant e^{-2}(1 - c_1 - c_2)|V_1|/O(\log k)$.

Now fix any sufficiently small $\epsilon > 0$. Letting $c_1 = 1/3 - \epsilon/2, c_2 = 2/3 - \epsilon$, we get that each of the first two cases imply that $f(G,1,k) \geqslant (2/3 - \epsilon - o_k(1))|V_1|/k \geqslant (1/3 - \epsilon - o_k(1))n/k$ (since $|V_1| \geqslant |V_2|$). Meanwhile with ϵ fixed, the third case implies $f(G,1,k) = \Omega_\epsilon(n/\log k)$. Taking $\epsilon \downarrow 0$ as $k \to \infty$ we have that $f(G,1,k) \geqslant (1/3 - o_k(1))n/k$. $\square$

[4]If $p = 0$, then $U = W_{k-1}$ and this probability is one. Otherwise this probability is at least $\binom{|N(v) \cap W_{k-1}|}{1}(1-2^{-p})^{|N(v)\cap W_{k-1}|}2^{-p} \geqslant (1-2^{-p})^{2^{p+1}-1} \geqslant e^{-2}$.

As a closing remark, we note it is still open whether $c(k) = 1/k$ for all k (as noted in [10], considering $K_{k,k}$ demonstrates that $c(k) \leqslant 1/k$). Even for $k = 2$, the best known bounds are $1/4 \leqslant c(2) \leqslant 1/2$, with the lower bound coming from [9, Theorem 2].

A An elementary proof of our lemma

Throughout, we let $\mathrm{Bin}(N, 1/2)$ denote the sum of N i.i.d. random variables that uniformly sample $\{0, 1\}$. For integers N, i, k, r, we let

$$p_{N,i} := \mathbb{P}(\mathrm{Bin}(N, 1/2) = i) = \binom{N}{i} 2^{-N},$$

$$P_{N,k,r} := \mathbb{P}(\mathrm{Bin}(N, 1/2) \equiv r \pmod{k}).$$

Also, for integer N we let $m_N := \max_i\{p_{N,i}\} = \binom{N}{\lfloor N/2 \rfloor} 2^{-N}$. A standard application of Stirling's formula gives $m_N \leqslant 1/\sqrt{N}$ for sufficiently large N (in fact $m_N\sqrt{N} \to \sqrt{2/\pi}$).

We will first show the following.

Proposition 4. *For all N, k, r, r', $|P_{N,k,r} - P_{N,k,r'}| \leqslant 2m_N$.*

Proof. First, we note that $|i - N/2| < |i' - N/2|$ implies $p_{N,i} > p_{N,i'}$ (since $\binom{N}{x}$ is a concave function of x on the domain $[0, N]$, which is symmetric and achieves it maximum at $x = N/2$). Hence, writing $I(E)$ to denote the indicator of an event E, we get

$$\begin{aligned} |P_{N,k,r} - P_{N,k,r'}| &\leqslant \left| \sum_{i=1}^{N/2} p_{N,i} (I(i \equiv r \pmod{k})) - I(i \equiv r' \pmod{k})) \right| \\ &\quad + \left| \sum_{i=1}^{N/2} p_{N,N-i} (I(N - i \equiv r \pmod{k})) - I(N - i \equiv r' \pmod{k})) \right| \\ &\leqslant 2m_N \end{aligned}$$

(here the last step follows from the fact that the non-zero summands of each sum alternate in sign, are increasing in absolute value, and are bounded in absolute value by m_N). $\square$

We can now establish the desired lemma.

Proof of Lemma 3. For integers N, k, let $M_{N,k} = \min_r\{P_{N,k,r}\}$.

Obviously, we always have $P_{N,k,1} \geqslant M_{N,k}$. So it suffices to show that for $N \geqslant k^3$, that $M_{N,k} \geqslant (1 - o_k(1))/k$.

By Proposition 4, we see that

$$1 = \sum_{r=1}^{k} P_{N,k,r} \leqslant k(M_{N,k} + 2m_N),$$

hence

$$M_{N,k} \leqslant (1 - 2km_N)/k.$$

Assuming $N \geqslant k^3$ with k sufficiently large, we have

$$M_{N,k} \leqslant (1 - 2/\sqrt{k})/k = (1 - o_k(1))/k$$

as desired. □

Acknowledgments

The author thanks Zachary Chase for spotting some typographical errors in a previous draft of this paper. We also thank the referees for some useful suggestions to improve the writing and references for this paper.

References

[1] P. Balister, E. Powierski, A. Scott, and J. Tan, *Counting partitions of $G(n, 1/2)$ with degree congruence conditions,* preprint (May 2021), arXiv:2105.12612.

[2] Y. Caro, *On induced subgraphs with odd degrees,* in *Discrete Mathematics,* **132** (1994), 23–28.

[3] Y. Caro, I. Krasikov and Y. Roditty, *Zero-sum partition theorems for graphs,* in *International Journal of Mathematics and Mathematical Sciences,* **17** (1994), 697–702.

[4] P. Diaconis, *Group representations in probability and statistics,* (1998).

[5] A. Ferber, L. Hiaman, M. Krivelevich, *On subgraphs with degrees of prescribed residues in the random graph,* preprint (July 2021), arXiv:2107.06977.

[6] A. Ferber and M. Krivelevich, *Every graph contains a linearly sized induced subgraph with all degrees odd,* in *Advances in Mathematics* (2022).

[7] L. Lovasz, *Combinatorial problems and exercises (2nd edition),* in *AMS Chelsea Publishing* (1993).

[8] J. Nagy and P. P. Pach, *A counterexample to the lights out problem,* in *Journal of Graph Theory* **101** (2022), 265–273.

[9] A. Scott, *Large induced subgraphs with all degrees odd,* in *Combinatorics, Probability and Computing* **1** (1992), 335–349.

[10] A. Scott, *On induced subgraphs with all degrees odd,* in *Graphs and Combinatorics* **17** (2001), 539–553.