

# Large induced subgraphs with all degrees odd

A.D. Scott

Department of Pure Mathematics and Mathematical Statistics,  
University of Cambridge, England

**Abstract:** We prove that every connected graph of order  $n \geq 2$  has an induced subgraph with all degrees odd of order at least  $cn/\log n$ , where  $c$  is a constant. We also give a bound in terms of chromatic number, and resolve the analogous problem for random graphs.

## 0. Introduction.

Given a graph  $G$ , it is natural to ask whether  $G$  must contain a large induced subgraph with certain properties. For instance, a classical question in Ramsey theory asks how many vertices a graph can have if it does not contain either a complete or an empty induced subgraph on a given number of vertices.

The starting point of this paper is the following theorem due to Gallai (see [4], Problem 5.17).

### Theorem A.

- (a) For every graph  $G$ , there is a partition  $V(G) = V_1 \cup V_2$  such that  $G[V_1]$  and  $G[V_2]$  have all degrees even.
- (b) For every graph  $G$ , there is a partition  $V(G) = V_1 \cup V_2$  such that  $G[V_1]$  has all degrees odd and  $G[V_2]$  has all degrees even. ■

As an immediate consequence of this theorem we have the following.

**Corollary B.** For every graph  $G$ , there is a set  $W \subset V(G)$  such that  $|W| \geq \frac{1}{2}|G|$  and  $G[W]$  has all degrees even. ■

Let us note that this result is best possible, as can be seen by considering  $P_n$ , the path on  $n$  vertices, for any  $n$ . (As we shall see below, we can say somewhat more for random graphs.)

It is natural to ask whether we can find analogous results for induced subgraphs with all degrees odd. If a graph  $G$  has an isolated vertex  $v$ , then no induced subgraph of  $G$  containing  $v$  can have all degrees odd. Thus, in order to ensure that  $G$  has a large subgraph of this form, it is natural to forbid isolated vertices. Equivalently, we may assume that  $G$  is connected: then if  $G$  has no isolated vertices, but is not connected, we may consider each component separately.

This leads to the conjecture (the origins of which are unclear, see [2]) that there is some  $\epsilon > 0$  such that every graph  $G$  without isolated vertices contains an induced subgraph of order at least  $\epsilon|G|$  with all degrees odd. If this were true then  $\epsilon$  could be at most  $\frac{1}{3}$ , as can be seen by considering  $K_6$  with a 1-factor removed.

In order to state the problem more clearly, we define some notation. For any graph  $G$  we define

$$f(G) = \max\{|W| : W \subset V(G) \text{ and } G[W] \text{ has all degrees odd}\}.$$

Then, for  $n \geq 2$ , let  $f(n) = \min\{f(G) : |G| = n \text{ and } \delta(G) \geq 1\}$ . The conjecture claims that there is some  $\epsilon > 0$  such that  $f(n) \geq \epsilon n$  for all  $n \geq 2$ .

The conjecture has been proved for trees: Caro, Krasikov and Roditty [3] gave a linear lower bound, and Radcliffe and Scott [5] proved a best possible bound, namely that  $f(T) \geq 2\lfloor(n+1)/3\rfloor$  for every tree  $T$ . In §1 we prove a bound in terms of chromatic number:  $f(G) \geq |G|/(2\chi(G))$  for every graph  $G$  without isolated vertices.

In the general case, Caro [2] proves that  $f(n) \geq \sqrt{(n - \sqrt{n})/6}$ . In §2 we prove the stronger result that  $f(n) \geq n/900 \log n$ . The result is generalised to weighted graphs in §3. Finally, in §4, we show that, whenever  $\omega(n) \rightarrow \infty$ , almost every  $G \in \mathcal{G}(n, \frac{1}{2})$  satisfies  $|f(G) - (cn + d \log n)| \leq \omega(n)$ , where  $c = 0.7729\dots$  and  $d = -0.2606\dots$ .

We use standard notation throughout (see [1], say). In particular, let  $G$  be a graph and  $S \subset V(G)$ . Then  $G[S]$  denotes the subgraph of  $G$  induced by  $S$ ; where unambiguous,

we write  $e(S)$  for  $e(G[S])$  and, for  $T \subset V(G) \setminus S$ , we write  $e(S, T)$  for  $|\{xy \in E(G) : x \in S, y \in T\}|$ . We will also write  $[n] = \{1, \dots, n\}$ .

### 1. A bound in terms of chromatic number.

We begin by giving a lower bound for  $f(G)$  in terms of the largest independent set in  $G$ . Alon (see [2]) has proved that  $f(G) \geq \frac{1}{3}\text{ind}(G)$ ; here we improve this to an essentially best possible result.

**Theorem 1.** *For any graph  $G$  without isolated vertices, we have  $f(G) \geq \frac{1}{2}\text{ind}(G)$ .*

**Proof.** Let  $G$  be a graph without isolated vertices, and let  $I \subset V(G)$  be an independent set of maximal size in  $G$ . We note that, for every  $v \in I$ , we have  $\Gamma(v) \cap (V(G) \setminus I) \neq \emptyset$ , since  $G$  contains no isolated vertices. Let  $W$  be a minimal subset of  $V(G) \setminus I$  such that  $|\Gamma(v) \cap W| > 0$  for every  $v \in I$ , say  $W = \{w_1, \dots, w_k\}$ . Then for each  $w_i$  we can find some  $u_i \in I$  such that  $\Gamma(u_i) \cap W = \{w_i\}$ . Set  $U = \{u_1, \dots, u_k\}$  and  $U' = I \setminus U$ .

For  $S \subset W$ , let  $U_S = \{u \in U' : |\Gamma(u) \cap S| \text{ is odd}\}$ , and

$$I_S = S \cup U_S \cup \{u_i : w_i \in S \text{ and } |\Gamma(w_i) \cap U_S| \text{ is even}\}.$$

By construction, in the graph  $G[I_S]$  every degree is odd, so  $f(G) \geq |I_S| \geq |S| + |U_S|$ .

Now let us choose  $S \subset W$  randomly, by picking each  $w_i$  independently, with probability  $\frac{1}{2}$ . Then, for every  $u \in U'$ ,  $\mathbb{P}(|\Gamma(u) \cap S| \text{ is odd}) = \frac{1}{2}$ , so

$$\begin{aligned} \mathbb{E}(|I_S|) &\geq \mathbb{E}(|S|) + \mathbb{E}(|U_S|) \\ &= |W|/2 + |U'|/2 \\ &= |I|/2, \end{aligned}$$

since  $|U'| = |I| - |U| = |I| - |W|$ . Therefore, for some  $S_0 \subset W$  we have  $|I_{S_0}| \geq |I|/2 = \frac{1}{2}\text{ind}(G)$ . ■

Theorem 1 is best possible in the sense that we cannot replace  $\frac{1}{2}$  with any smaller constant. Indeed, let  $B_n$  be the bipartite graph with vertex classes  $V_1 = [n]$ ,  $V_2 = [n]^{(2)}$ ,

and edges from  $\{i\}$  to  $\{j, k\}$  if and only if  $i \in \{j, k\}$ . We note that if  $W \subset V(B_n)$ , then for each  $\{i, j\} \in W \cap V_2$  exactly one of  $\{i\}$  and  $\{j\}$  are in  $V_1$ . Thus  $f(B_n) \leq \lceil n^2/4 \rceil + n$ . Since  $\text{ind}(B_n) = \binom{n}{2}$ , we have  $f(B_n)/\text{ind}(B_n) \rightarrow \frac{1}{2}$  as  $n \rightarrow \infty$ , and deduce that the constant in Theorem 1 is best possible.

As a consequence of Theorem 1 we have the following result.

**Theorem 2.** *Let  $G$  be a graph without isolated vertices. Then  $f(G) \geq |G|/(2\chi(G))$ .*

**Proof.** Colour  $G$  with  $\chi(G)$  colours and apply Theorem 1 to the largest colour class. ■

How good is the factor  $(2\chi(G))^{-1}$  in Theorem 2? For  $k \geq 2$ , we define

$$c_k = \inf\{f(G)/|G| : \chi(G) = k \text{ and } \delta(G) \geq 1\},$$

Thus Theorem 2 says that  $c_k \geq 1/2k$ . For bipartite graphs, this gives  $c_2 \geq \frac{1}{4}$ ; from the other side, the graph  $B_n$ , defined above, shows that  $c_2 \leq \frac{1}{2}$ , and we conjecture that  $\frac{1}{2}$  is the correct value. For  $k \geq 3$ , our best upper bound is  $c_k \leq \frac{1}{3}$ , which we get by considering  $K_6$  with a 1-factor removed (this graph is 3-colourable). It seems likely that some improvement is possible.

The  $c_k$  clearly form a non-increasing sequence (indeed, suppose  $k \leq l$  and  $G$  is  $k$ -chromatic: then by considering many disjoint copies of  $G$  together with a copy of  $K_{l+1}$ , we see  $f(G) \geq c_l|G|$ ; we deduce  $c_k \geq c_l$ ). The remainder of this section examines how quickly the sequence can decrease.

Suppose  $k \geq l$ , and we wish to bound  $c_k$  below in terms of  $c_l$ . Let  $G$  be a  $k$ -chromatic graph. Given a  $k$ -colouring of  $G$ , we would like to take the  $l$  largest colour classes, say  $V_1, \dots, V_l$ , and by considering  $G[\bigcup_{i=1}^l V_i]$  and our lower bound for  $c_l$  obtain a lower bound for  $f(G)$ . Unfortunately, although  $G[\bigcup_{i=1}^l V_i]$  is  $l$ -chromatic, it may contain isolated vertices. This leads us to ask whether  $G$  must contain a large induced  $l$ -chromatic graph without isolated vertices.

In order to answer this question, it seems to be easiest to prove a more general result concerning weighted graphs. We say that a graph  $G$  has *vertex weighting*  $\mu$  if  $\mu$  is a function  $\mu: V(G) \rightarrow \mathbb{R}_{\geq 0}$ . The *weight* of a vertex  $v \in V(G)$  is  $\mu(v)$ , and the *weight* of a

set  $W \subset V(G)$  is  $\mu(W) = \sum_{w \in W} \mu(w)$ . If  $H$  is a subgraph of  $G$  then the *weight* of  $H$  is  $\mu(H) = \mu(V(H))$ . We have the following result for vertex-weighted graphs.

**Theorem 3.** *Suppose  $k \geq 2$  and  $G$  is a  $k$ -chromatic graph with vertex-weighting  $\mu$  and no isolated vertices. Then, for each  $2 \leq m \leq k$ ,  $G$  contains an induced  $m$ -chromatic subgraph  $H$  such that  $H$  has no isolated vertices and*

$$\mu(H) \geq \frac{m}{k} \mu(G).$$

**Proof.** If  $k = m = 2$  then the theorem is trivial. Let us suppose that  $k \geq 3$ . We will prove by induction on  $|V(G)|$  that if  $G$  is a  $k$ -chromatic graph with vertex-weighting  $\mu$  and no isolated vertices then  $G$  contains an induced  $(k - 1)$ -chromatic subgraph with weight at least  $\frac{k-1}{k} \mu(G)$  and without isolated vertices. Applying this  $k - m$  times gives the required  $m$ -chromatic subgraph.

If  $|V(G)| = 3$  then the inductive statement is trivial. Let us assume that  $|V(G)| > 3$ , that  $G$  is  $k$ -chromatic and has no isolated vertices, and that our inductive statement is true for smaller values of  $|V(G)|$ . We may assume that  $G$  is connected, or else deal with each component of  $G$  separately. Let  $V(G) = V_1 \cup \dots \cup V_k$  be a  $k$ -colouring of the vertices of  $G$ .

Suppose that some  $v \in V(G)$  has all its neighbours in one colour class, say  $v \in V_1$  and  $\Gamma(v) \subset V_2$ . Then consider  $G' = (G \setminus (\{v\} \cup \Gamma(v))) \cup \{w\}$ , where  $w$  is a new vertex joined to  $\bigcup_{x \in \Gamma(v)} \Gamma(x) \setminus \{v\}$ . Note that  $w$  is not an isolated vertex in  $G'$ , since otherwise  $\{v\} \cup \Gamma(v)$  would have been a component of  $G$  with chromatic number 2. Let  $V'_1 = V_1 \setminus \{v\}$ ,  $V'_2 = (V_2 \setminus \Gamma(v)) \cup \{w\}$  and  $V'_i = V_i$  for  $i > 2$ . Then  $G'$  is  $k$ -chromatic and  $V'_1, \dots, V'_k$  is a  $k$ -colouring. Define  $\mu'(x) = \mu(x)$  for  $x \in V(G') \setminus \{w\}$  and  $\mu'(w) = \mu(v) + \sum_{x \in \Gamma(v)} \mu(x)$ . Thus  $\mu'(G') = \mu(G)$ . Since  $|V(G')| < |V(G)|$ , we can find a  $(k - 1)$ -chromatic induced subgraph  $H' \subset G'$  without isolated vertices and with

$$\mu'(H') \geq \frac{k-1}{k} \mu'(G') = \frac{k-1}{k} \mu(G).$$

If  $w \in V(H')$  then let  $H = G[(V(H') \setminus \{w\}) \cup \{v\} \cup \Gamma(v)]$ ; if  $w \notin V(H')$  then let  $H = H'$  (considered as a subgraph of  $G$ ). We see that  $\mu(H) = \mu'(H')$ , and  $H$  is an induced  $(k - 1)$ -chromatic subgraph of  $G$  without isolated vertices.

Thus we may assume that every vertex in  $G$  has neighbours in at least two colour classes. In that case, we may delete the colour class of smallest weight, to obtain the required induced subgraph of  $G$ . ■

This gives an immediate corollary for unweighted graphs.

**Corollary 4.** *Let  $2 \leq m \leq k$ . If  $G$  is a  $k$ -chromatic graph without isolated vertices then it contains an induced  $m$ -chromatic graph  $H$  without isolated vertices such that*

$$|H| \geq \frac{m}{k}|G|.$$

**Proof.** Define  $\mu(v) = 1$  for all  $v \in V(G)$  and apply Theorem 3. ■

Both Theorem 3 and Corollary 4 are best possible, as can be seen by considering a complete  $k$ -partite graph with all vertex classes of equal size.

We are now ready to prove our bound on how fast  $\{c_k\}_{k=2}^\infty$  can decrease.

**Theorem 5.** *For every  $2 \leq k \leq l$ ,  $c_l \geq \frac{k}{l}c_k$ .*

**Proof.** Let  $G$  be an  $l$ -chromatic graph. Then by Corollary 4, we can find an induced  $k$ -chromatic subgraph  $H$  such that  $H$  has no isolated vertices and  $|H| \geq k|G|/l$ . Therefore

$$f(G) \geq f(H) \geq c_k|H| \geq \frac{kc_k}{l}|G|.$$

We deduce that  $c_l \geq kc_k/l$ . ■

We note that any improvement in the lower bound for  $c_k$  gives corresponding improvements in the lower bounds for  $c_l$ , for every  $l \geq k$ . In particular, it would be interesting to improve  $c_2$ , since this would lead to a corresponding improvement on the factor  $(2\chi(G))^{-1}$  in Theorem 2.

It is not impossible that the conjecture  $f(n) \geq cn$  for all  $n \geq 2$ , for some constant  $c$ , is false. Even if it is true, we may only be able to prove it with  $c$  small, in which case Theorem 2 will give a better bound for graphs of low chromatic number. If  $c_2 = \frac{1}{2}$ , as conjectured above, an application of Theorem 5 would improve Theorem 2 by a factor of

2: we would get the attractive result that  $f(G) \geq |G|/\chi(G)$  for every graph  $G$  without isolated vertices.

## 2. A general bound.

We have proved a bound in terms of chromatic number; we now turn to the general case. The idea of the proof is to look for subgraphs of  $G$  satisfying conditions which ensure that  $G$  contains relatively large induced subgraphs with all degrees odd.

For instance, suppose we have an edge  $vw \in E(G)$  with  $d(v)$  large and  $|\Gamma(v) \cap \Gamma(w)|$  small. Let  $S = \Gamma(v) \setminus (\Gamma(w) \cup \{w\})$ . Then, by Theorem B, we can find  $S' \subset S$  such that  $|S'| \geq |S|/2$  and  $G[S']$  has all degrees even. If  $|S|$  is even then set  $W = S' \cup \{v, w\}$ ; if  $|S|$  is odd, then set  $W = S' \cup \{v\}$ . In either case,  $G[W]$  has all degrees odd and  $|W| \geq \frac{1}{2}(d(v) - |\Gamma(v) \cap \Gamma(w)|)$ .

More generally, suppose we have a set of independent edges  $v_1w_1, \dots, v_kw_k$ , and let  $V_0 = \bigcup_{i=1}^k \Gamma(v_i)$  and  $W_0 = \bigcup_{i=1}^k \Gamma(w_i)$ . If  $|V_0 \cap W_0|$  is small compared with  $|V_0 \cup W_0|$ , then we can use a similar argument to find a comparatively large induced subgraph of  $G$  with all degrees odd. The main part of the proof works by looking for such collections of edges.

**Theorem 6.** *Every graph  $G$  without isolated vertices contains a set  $W \subset V(G)$  such that  $|W| \geq |G|/900 \log(|G|)$  and  $G[W]$  has all degrees odd.*

**Proof.** Let  $G$  be a graph of order  $n$  without isolated vertices, and let  $c$  be a constant (we will take  $c = 0.086$  below). We begin by partitioning  $V(G)$  into sets  $X_1, Y_1$  and  $Z_1$ . We will do this in such a way that  $X_1$  contains a large structure of the sort described above;  $Y_1$  has highly connected components; and  $Z_1$  does not contain too many edges. Both  $X_1$  and  $Y_1$  can be used to find relatively large induced subgraphs of  $G$  with all degrees odd; our only problem is  $Z_1$ , which we again partition, into sets  $X_2, Y_2$  and  $Z_2$ . We apply the algorithm repeatedly, each time partitioning  $Z_{i-1}$  into sets  $X_i, Y_i$  and  $Z_i$ , ending up with a partition of  $V(G)$  into sets  $X_1, \dots, X_k, Y_1, \dots, Y_k, Z_k$ , where  $Z_k$  is an independent set.

More precisely, let  $Z_0 = V(G)$ . By assumption,  $e(Z_0) > 0$ : we partition  $Z_0$  into

sets  $X_1, Y_1$  and  $Z_1$  using the algorithm described below. If  $e(Z_1) = 0$  then set  $k = 1$ . Otherwise, apply the algorithm to  $Z_1$  to get  $X_2, Y_2$  and  $Z_2$ . Repeat this process until  $Z_i$  is first an independent set, at which point put  $k = i$ .

Our algorithm for partitioning  $Z_{j-1}$  into  $X_j, Y_j$  and  $Z_j$  is as follows. We shall use the notation  $X_j^{(i)}, Y_j^{(i)}$  and  $Z_j^{(i)}$  for our successive ‘approximations’ to  $X_j, Y_j$  and  $Z_j$ . Let  $X_j^{(1)} = Y_j^{(1)} = Z_j^{(1)} = \emptyset$ . At the  $i$ th stage, given  $X_j^{(i)}, Y_j^{(i)}$  and  $Z_j^{(i)}$ , we define

$$V_j^{(i)} = Z_{j-1} \setminus (X_j^{(i)} \cup Y_j^{(i)} \cup Z_j^{(i)}).$$

Thus  $V_j^{(i)}$  is the set of vertices we have not yet partitioned. If there is some  $v \in V_j^{(i)}$  which satisfies

$$|\Gamma(v) \cap X_j^{(i)}| \geq c|\Gamma(v) \cap V_j^{(i)}| \quad (1)$$

then set

$$\begin{aligned} X_j^{(i+1)} &= X_j^{(i)} \\ Y_j^{(i+1)} &= Y_j^{(i)} \\ Z_j^{(i+1)} &= Z_j^{(i)} \cup \{v\}. \end{aligned}$$

Otherwise, if there is an edge  $vw$  in  $G[V_j^{(i)}]$  such that  $|\Gamma(v) \cap V_j^{(i)}| > 1$  and

$$|\Gamma(v) \cap \Gamma(w) \cap V_j^{(i)}| \leq (2/3) \max\{|\Gamma(v) \cap V_j^{(i)}| - 1, |\Gamma(w) \cap V_j^{(i)}| - 1\}, \quad (2)$$

then set

$$\begin{aligned} X_j^{(i+1)} &= X_j^{(i)} \cup ((\Gamma(v) \cup \Gamma(w)) \cap V_j^{(i)}) \setminus \{v, w\} \\ Y_j^{(i+1)} &= Y_j^{(i)} \cup \{v, w\} \\ Z_j^{(i+1)} &= Z_j^{(i)}. \end{aligned} \quad (3)$$

If neither of these conditions is satisfied, then set

$$\begin{aligned} X_j &= X_j^{(i)} \\ Y_j &= Y_j^{(i)} \\ Z_j &= Z_j^{(i)}, \end{aligned} \quad (4)$$



and terminate the algorithm.

We have now partitioned  $V(G)$  into  $2k + 1$  sets. Note that the number  $e(Z_j)$  of edges in  $G[Z_j]$  satisfies  $e(Z_j) \leq e(Z_{j-1})/(1 + c)$ , for  $j = 1, \dots, k$ . Indeed, let  $Z_i = \{z_1, \dots, z_p\}$ , where  $z_j$  is chosen at stage  $i_j$  and  $i_1 < \dots < i_p$ . Using (1), we have

$$\begin{aligned} e(Z_i) &= \sum_{j=1}^{p-1} |\Gamma(z_j) \cap \{z_{j+1}, \dots, z_p\}| \\ &\leq \sum_{j=1}^{p-1} |\Gamma(z_j) \cap V_i^{(i_j)}| \\ &\leq \sum_{j=1}^{p-1} |\Gamma(z_j) \cap X_i^{(i_j)}|/c \\ &\leq e(Z_i, X_i)/c. \end{aligned}$$

Since  $X_i \cup Z_i \subset Z_{i-1}$ , we have

$$e(Z_{i-1}) \geq e(Z_i) + e(Z_i, X_i) \geq (1 + c)e(Z_i),$$

as claimed. Thus  $k$  (the smallest integer such that  $Z_k$  is an independent set) is well defined, and

$$k \leq \log_{1+c} \binom{n}{2} \leq 2 \log n / \log(1 + c). \quad (5)$$

We claim that each set in our partition guarantees the existence of a comparatively large induced subgraph of  $G$  with all degrees odd. The next three lemmas prove that this is indeed the case.

**Lemma 7.** For  $i = 1, \dots, k$ , we have  $f(G) \geq \frac{1-6c}{16}|X_i|$ .

**Proof.** If  $X_i = \emptyset$  then we are done. Otherwise,  $X_i$  is in the neighbourhood of a sequence of edges  $v_1w_1, \dots, v_pw_p$  chosen to satisfy (2). Let us suppose that  $v_jw_j$  was chosen at the  $i_j$ th stage, where  $i_1 < \dots < i_p$ , and that

$$|\Gamma(v_j) \cap V_i^{(i_j)}| \geq |\Gamma(w_j) \cap V_i^{(i_j)}|. \quad (6)$$

Let  $V_0 = \{v_1, \dots, v_p\}$ ,  $V_1 = \{w_1, \dots, w_p\}$  and  $S = \{v \in X_i : \Gamma(v) \cap V_1 = \emptyset\}$ . We remark that, from (2) and (3),  $V_0$  and  $V_1$  are independent sets of  $Y_i$ , that  $\Gamma(v_i) \cap V_1 = \{w_i\}$ ,

and that every  $v \in S$  has at least one neighbour in  $V_0$ . Let  $T \subset V_0$  be a random subset, where each  $v \in V_0$  is chosen independently with probability  $\frac{1}{2}$ . We define

$$S_T = \{v \in S : |\Gamma(v) \cap T| \text{ is odd}\}.$$

Clearly,  $\mathbb{E}(|S_T|) = |S|/2$ , so for some  $T_0 \subset V_0$  we have  $|S_{T_0}| \geq |S|/2$ . By Corollary B, we can find some  $S_0 \subset S_{T_0}$  such that  $|S_0| \geq |S_{T_0}|/2 \geq |S|/4$  and  $G[S_0]$  has all degrees even. Let

$$S' = S_0 \cup T_0 \cup \{w_i \in V_1 : v_i \in T_0 \text{ and } |\Gamma(v_i) \cap S_0| \text{ is even}\}.$$

It is easily seen that  $G[S']$  has all degrees odd and  $|S'| \geq |S_0| \geq |S|/4$ .

In order to prove the assertion of the lemma, it is therefore enough to prove that  $|S| \geq (1 - 6c)|X_i|/4$ . We do this by using inequalities (2) and (6), and following through the algorithm for partitioning  $Z_{i-1}$ . This will require some calculation which appears rather unattractive, but is in fact quite straightforward.

Define  $S_0 = \emptyset$  and, for  $j = 1, \dots, p$ ,

$$S_j = \{v \in X_i^{(i_j)} : \Gamma(v) \cap \{w_1, \dots, w_j\} = \emptyset\}, \quad (7)$$

Note that  $S_p = S$  and  $X_i^{(i_p)} = X_i$ , so we want to show that  $|S_p| \geq (1 - 6c)|X_i^{(i_p)}|/4$ . Set  $i_0 = 0$ , so that  $X_i^{(i_0)} = X_i^{(0)} = \emptyset$ . From (3) and (7) we see that, for  $j = 1, \dots, p$ ,

$$S_j = (S_{j-1} \setminus \Gamma(w_j)) \cup (\Gamma(v_j) \cap V_i^{(i_j)} \setminus (\{w_j\} \cup \Gamma(w_j))),$$

so

$$|S_j| \geq |S_{j-1}| - |\Gamma(w_j) \cap S_{j-1}| + |\Gamma(v_j) \cap V_i^{(i_j)}| - 1 - |\Gamma(v_j) \cap \Gamma(w_j) \cap V_i^{(i_j)}|. \quad (8)$$

On the other hand, from (3), we have

$$X_i^{(i_j)} = X_i^{(i_{j-1})} \cup ((\Gamma(v_j) \cup \Gamma(w_j)) \cap V_i^{(i_j)} \setminus \{v, w\}),$$

so

$$\begin{aligned} |X_i^{(i_j)}| - |X_i^{(i_{j-1})}| &= |\Gamma(v_j) \cap V_i^{(i_j)}| + |\Gamma(w_j) \cap V_i^{(i_j)}| - |\Gamma(v_j) \cap \Gamma(w_j) \cap V_i^{(i_j)}| - 2 \\ &\leq 2(|\Gamma(v_j) \cap V_i^{(i_j)}| - 1) - |\Gamma(v_j) \cap \Gamma(w_j) \cap V_i^{(i_j)}|, \end{aligned} \quad (9)$$

by inequality (6). For ease of notation we write  $a = |\Gamma(v_j) \cap V_i^{(i_j)}|$  and  $b = |\Gamma(v_j) \cap \Gamma(w_j) \cap V_i^{(i_j)}|$ . Note that, from (1),  $|\Gamma(w_j) \cap S_{j-1}| \leq |\Gamma(w_j) \cap X_i^{(i_j)}| \leq c|\Gamma(w_j) \cap V_i^{(i_j)}| \leq c|\Gamma(v_j) \cap V_i^{(i_j)}| = ca$ . Then (8) and (9) become

$$|S_j| - |S_{j-1}| \geq (a - 1) - b - ca,$$

and

$$|X_i^{(i_j)}| - |X_i^{(i_{j-1})}| \leq 2(a - 1) - b.$$

Inequality (2) becomes  $b \leq 2(a - 1)/3$ , so

$$\begin{aligned} \frac{|S_j| - |S_{j-1}|}{|X_i^{(i_j)}| - |X_i^{(i_{j-1})}|} &\geq \frac{(a - 1) - b - ca}{2(a - 1) - b} \\ &\geq \frac{\frac{1}{3}(a - 1) - ca}{\frac{4}{3}(a - 1)} \\ &= \frac{1 - 3c}{4} - \frac{3c}{4(a - 1)} \\ &\geq \frac{1 - 6c}{4}, \end{aligned}$$

since  $a > 1$ . This is true for  $j = 1, \dots, p$ , so

$$\begin{aligned} |S_p| &= \sum_{j=1}^p (|S_j| - |S_{j-1}|) \\ &\geq \sum_{j=1}^p \frac{1 - 6c}{4} (|X_i^{(i_j)}| - |X_i^{(i_{j-1})}|) \\ &= \frac{1 - 6c}{4} |X_p^{(i_p)}| \end{aligned}$$

.

■

**Lemma 8.** For  $i = 1, \dots, k$ ,  $f(G) \geq \frac{1}{4}|Y_i|$ .

**Proof.** We note first that  $G[Y_i]$  has no isolated vertices. Indeed, every vertex  $v \in Y_i$  was added in (3) or (4). If  $v$  was added in (3), then it is adjacent to some  $w \in Y_i$  that was added at the same time; while if  $v$  was added in (4), then  $|\Gamma(v) \cap Y_i| \geq |\Gamma(v) \cap V_j^{(i)}| > 0$ , or else  $v$  would have satisfied the condition (1).

We consider the components of  $G[Y_i]$ . The vertices added to  $Y_i$  in (3) induce a set of independent edges, so they are all in components of order 2; components of great order must come from (4). We claim that each component  $H$  contains a set  $V_H \subset V(H)$  such that  $|V_H| \geq \frac{1}{4}|H|$  and  $G[V_H]$  has all degrees odd. This is enough to prove the lemma, since the union of these sets has order at least  $|Y_i|/4$  and induces a subgraph with all degrees odd. We will suppress unnecessary mention of  $H$ , writing for instance  $\Gamma$  for  $\Gamma_H$ .

If  $|H| = 2$  then we are done. Otherwise,  $|H| > 2$ , and  $V(H)$  was added to  $Y_i$  in (4). By (2), each adjacent pair of vertices  $x$  and  $y$  in  $H$  satisfy

$$|\Gamma(x) \cap \Gamma(y)| > \frac{2}{3} \max\{d(x) - 1, d(y) - 1\}. \quad (10)$$

Note that if  $xy \in E(H)$  and  $d(x) \geq 2$  then  $d(y) \geq 2$ ; since  $H$  is connected, we may assume that  $\delta(H) \geq 2$ .

Suppose that  $\Delta(H) \geq \frac{1}{2}|H|$ , and let  $x \in V(H)$  be a vertex of maximal degree in  $H$ . If  $d(x)$  is odd then let  $W = \Gamma(x)$ ; otherwise let  $W = \Gamma(x) \setminus \{y\}$  for some  $y \in \Gamma(x)$ . By Theorem A we may split  $V(W) \cup \{x\}$  into sets  $W_1$  and  $W_2$ , each of which induces a subgraph of  $H$  with all degrees even. We may assume that  $x \in W_1$ . Let  $W'_1 = W_1 \setminus \{x\}$ ,  $W'_2 = W_2 \cup \{x\}$ . The subgraph  $W'_1$  has all degrees odd, and thus an even number of vertices.  $G[W'_2]$  has all degrees odd, except possibly for  $x$ ; but  $|\Gamma(x) \cap W'_2| = |W_2| = |W| + 1 - |W_1| = |W| - |W'_1|$ , which is odd, so  $x$  has an odd number of neighbours in  $W'_2$ . Since  $|W'_1| + |W'_2| \geq \Delta(H) \geq \frac{1}{2}|H|$ , one of  $W'_1$  and  $W'_2$  must contain at least  $\frac{1}{4}|H|$  vertices, and will thus do for  $V_H$ .

It is therefore enough to prove that if  $H$  satisfies (10) then  $\Delta(H) \geq \frac{1}{2}|H|$ .

We claim first that  $\text{diam}(H) \leq 2$ . Indeed, suppose  $\text{diam}(H) > 2$ , say  $d(w, z) = 3$  for some  $w, z \in H$ . Let  $wxyz$  be a shortest path from  $w$  to  $z$ . We will prove that  $|\Gamma(w) \cap \Gamma(z)| > 0$ .

We may assume  $d(x) \geq d(y)$ . Now  $\Gamma(w) \cap \Gamma(x)$  and  $\Gamma(y) \cap \Gamma(x)$  are subsets of  $\Gamma(x) \setminus \{w, y\}$ , so by (10),

$$\begin{aligned} |\Gamma(w) \cap \Gamma(y)| &\geq |\Gamma(w) \cap \Gamma(x)| + |\Gamma(y) \cap \Gamma(x)| - (|\Gamma(x)| - 2) \\ &> \frac{2}{3}(d(x) - 1) + \frac{2}{3}(d(x) - 1) - d(x) + 2 \\ &= \frac{1}{3}d(x) + \frac{2}{3}. \end{aligned}$$

Since  $d(x) \geq d(y)$  we have

$$|\Gamma(w) \cap \Gamma(y)| \geq \frac{1}{3}d(y) + \frac{2}{3}.$$

We get

$$\begin{aligned} |\Gamma(w) \cap \Gamma(z)| &\geq |\Gamma(w) \cap \Gamma(y)| + |\Gamma(z) \cap \Gamma(y)| - |\Gamma(y)| \\ &> \frac{1}{3}d(y) + \frac{2}{3} + \frac{2}{3}(d(y) - 1) - d(y) \\ &= 0. \end{aligned}$$

Thus we must have  $d(w, z) \leq 2$ , which is a contradiction. Hence  $\text{diam}(H) \leq 2$ .

We now prove that  $\Delta(H) \geq \frac{1}{2}|H|$ . Indeed, suppose  $d(x) = \Delta(H) < \frac{1}{2}|H|$ . Let  $Y = \Gamma(x)$  and  $Z = H \setminus (Y \cup \{x\})$ . Then  $Y$  and  $Z$  are nonempty. We count  $e(Y, Z)$  in two ways.

First, for  $y \in Y$ , we have by (10),

$$\begin{aligned} e(y, Z) &= |\Gamma(y)| - |\Gamma(x) \cap \Gamma(y)| - 1 \\ &< d(y) - \frac{2}{3}(d(y) - 1) - 1 \\ &= \frac{1}{3}(d(y) - 1) \\ &\leq \frac{1}{3}(d(x) - 1). \end{aligned} \tag{11}$$

For  $z \in Z$ , pick  $w \in \Gamma(z) \cap \Gamma(x)$  (we can do this because  $\text{diam}(H) \leq 2$ ). We have, as before,

$$\begin{aligned} e(z, Y) &= |\Gamma(z) \cap \Gamma(x)| \\ &\geq |\Gamma(z) \cap \Gamma(w)| + |\Gamma(x) \cap \Gamma(w)| - (\Gamma(w) - 2) \\ &> \frac{2}{3}(d(w) - 1) + \frac{2}{3}(d(x) - 1) - (d(w) - 2) \\ &= \frac{2}{3}d(x) - \frac{1}{3}d(w) + \frac{2}{3} \\ &\geq \frac{1}{3}d(x) + \frac{2}{3}. \end{aligned} \tag{12}$$

Therefore, using (11) and (12) to bound  $e(Y, Z)$  above and below, we get

$$\frac{1}{3}(d(x) - 1)|Y| > e(Y, Z) > \frac{1}{3}(d(x) + 2)|Z|.$$

If  $d(x) < |H|/2$  then  $|Z| > d(x) - 1$ , so

$$\frac{1}{3}(d(x) - 1)d(x) > \frac{1}{3}(d(x) + 2)(d(x) - 1),$$

which is a contradiction. Thus  $\Delta(H) \geq \frac{1}{2}|H|$ . ■

**Lemma 9.** *We have  $f(G) \geq |Z_k|/2$ .*

**Proof.** Since  $Z_k$  is an independent subset of  $V(G)$ , this is immediate from Theorem 1. ■

We are now ready to complete the proof of Theorem 6. From the preceding lemmas we have

$$\begin{aligned} f(G) &\geq \max \left\{ \frac{1-6c}{16} \max_{i=1,\dots,k} |X_i|, \frac{1}{4} \max_{i=1,\dots,k} |Y_i|, \frac{1}{2}|Z_k| \right\} \\ &\geq \max \left\{ \frac{1-6c}{16k} \sum_{i=1}^k |X_i|, \frac{1}{4k} \sum_{i=1}^k |Y_i|, \frac{1}{2}|Z_k| \right\}. \end{aligned}$$

Now  $|G| = \sum_{i=1}^k |X_i| + \sum_{i=1}^k |Y_i| + |Z_k|$ , so by (5) we have

$$\begin{aligned} f(G) &\geq \frac{n}{16k/(1-6c) + 4k + 2} \\ &\geq \left( 8 \left( \frac{5-6c}{1-6c} \right) \frac{\log n}{\log(1+c)} + 2 \right)^{-1} n. \end{aligned}$$

Set  $c = 0.086$ . Since  $f(G) \geq 2$  always, we get  $f(G) \geq n/900 \log n$ . ■

If  $G$  is triangle-free then the proof of Theorem 6 yields a better constant in our bound.

**Corollary 10.** *Let  $G$  be a triangle-free graph without isolated vertices. Then  $f(G) \geq |G|/155 \log |G|$ .*

**Proof.** Let  $n = |G|$ . We use the notation of the proof of Theorem 6. Note that, since  $G$  is triangle-free, (2) is satisfied for every edge in  $G[V_j^{(i_j)}]$ . Since  $\Gamma(v) \cap \Gamma(w) = \emptyset$  whenever  $v$  and  $w$  are adjacent, we have  $b = |\Gamma(v_j) \cap \Gamma(w_j) \cap V_i^{(i_j)}| = 0$  in Lemma 7, so

$$\begin{aligned} \frac{|S_j| - |S_{j-1}|}{|X_i^{(i_j)}| - |X_i^{(i_{j-1})}|} &\geq \frac{(a-1) - ca}{2(a-1)} \\ &= \frac{1}{2} - \frac{ca}{2(a-1)} \\ &\geq \frac{1}{2} - c, \end{aligned}$$

since  $a > 1$ . It follows from (3) and (4) that  $G[Y_i]$  is a collection of independent edges, so Lemma 8 becomes  $f(G) \geq |Y_i|$ . Therefore

$$\begin{aligned} f(G) &\geq \max \left\{ \frac{1-2c}{8k} \sum_{i=1}^k |X_i|, \frac{1}{k} \sum_{i=1}^k |Y_i|, \frac{1}{2} |Z_k| \right\} \\ &\geq \left( \frac{8k}{1-2c} + k + 2 \right)^{-1} n \\ &\geq \left( \frac{9-2c}{1-2c} \frac{2 \log n}{\log(1+c)} + 2 \right)^{-1} n. \end{aligned}$$

Set  $c = 0.24$ : we get that  $f(G) > n/155 \log n$ . ■

We note that any restriction on  $\text{cl}(G)$  gives a corresponding restriction on the size of components in  $Y_i$  and thus a small improvement in the constant in our bound for  $f(G)$ . (For instance, if  $\text{cl}(G) = 3$  then  $Y_i$  can contain only triangles and independent edges.)

### 3. Induced subgraphs of weighted graphs.

We now consider weighted versions of our results. For a graph  $G$  with vertex-weighting  $\mu$ , let  $f_\mu(G)$  be the maximum weight of an induced subgraph of  $G$  with all degrees odd. We have the following equivalent to Theorem 1.

**Theorem 11.** *Let  $G$  be a graph with vertex-weighting  $\mu$  and without isolated vertices, and  $I$  be an independent subset of  $V(G)$ . Then*

$$f_\mu(G) \geq \frac{1}{4} \mu(I).$$

**Proof.** As in Theorem 1, pick  $W \subset V \setminus I$  to be a minimal set covering  $I$  and find  $U \subset I$  so that  $G[V \cup W]$  is a set of independent edges. If  $\mu(U) \geq \mu(I)/2$  then partition  $W$  into  $W_1$  and  $W_2$  so that  $G[W_1]$  and  $G[W_2]$  have all degrees even. One of  $W_1 \cup (U \cap \Gamma(W_1))$  and  $W_2 \cup (U \cap \Gamma(W_2))$  will do. If  $\mu(U) < \mu(I)/2$ , choose  $S \subset I \setminus U$  as in Theorem 1. Then  $\mathbb{E}(\mu(U_S)) = \mu(I \setminus U)/2 > \mu(I)/4$ , so we may proceed as before. ■

This can be used in the same way that Theorem 1 was used for Theorem 6.

**Theorem 12.** *There is a constant  $c$  such that for every graph  $G$  with vertex-weighting  $\mu$  and without isolated vertices,*

$$f_\mu(G) \geq c\mu(G)/\log |G|.$$

**Proof.** (Sketch) We follow closely the proof of Theorem 6, replacing  $|\cdot|$  with  $\mu(\cdot)$  where appropriate. We replace (1) with

$$\mu(\Gamma(v) \cap X_j^{(i)}) > c\mu(\Gamma(v) \cap V_j^{(i)}),$$

and (2) with the conditions

$$\mu(\Gamma(v) \cap V_j^{(i)} \setminus \{w\}) > 0$$

and

$$\mu(\Gamma(v) \cap \Gamma(w) \cap V_j^{(i)}) \leq (2/3) \max\{\mu(\Gamma(v) \cap V_j^{(i)}) - \mu(w), \mu(\Gamma(w) \cap V_j^{(i)}) - \mu(v)\}.$$

In the final step of the algorithm we leave the definitions of  $X_j$ ,  $Y_j$  and  $Z_j$  unchanged. ■

We can use this to find large induced subgraphs with all degrees odd that satisfy additional conditions on the distribution of vertices. For instance, we have the following more general version of Theorem 6.

**Theorem 13.** *There is a constant  $c$  such that for every graph  $G$  without isolated vertices and for every  $S \subset V(G)$  there is an induced subgraph  $H$  of  $G$  with all degrees odd such that*

$$|V(H) \cap S| \geq c|S|/\log |G|.$$

**Proof.** Set  $\mu = \chi_S$ , the indicator function for  $S$ , and apply Theorem 12. ■

We note that Theorem 13 can also be proved directly from Theorem 6 and Theorem 1: if  $\text{ind}(G[S]) > |G|/2$  then apply Theorem 1; otherwise apply Theorem 6 to each non-trivial component of  $G[S]$ .



#### 4. Random graphs.

So far we have considered results that hold for all graphs. In this section, we consider the same questions for random graphs.

Caro [2] has proved that almost every graph (in the model  $\mathcal{G}(n, 1/2)$ ) satisfies  $f(G) \geq n/4$ , and Alon (see [2]) has proved that almost every graph satisfies  $f(G) \geq (\frac{1}{2} - \epsilon)|G|$ . We prove a more precise result.

Let  $s(n)$  be the smallest positive even integer such that  $2^{1-s} \binom{n}{s} \leq 1$ , and let  $c$  be the solution of  $(2c)^c(1-c)^{1-c} = 1$ . Then  $c = 0.7729\dots$  and  $s(n) = cn + d \log n + O(1)$ , where  $1/d = 2 \log(\frac{1-c}{2c}) = -3.8359\dots$ .

**Theorem 14.** *Let  $s(n)$  be as above, and let  $\omega(n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Then almost every graph  $G \in \mathcal{G}(n, 1/2)$  satisfies*

$$|f(G) - s(n)| \leq \omega(n).$$

*In particular, almost every graph satisfies*

$$f(G) \geq 0.7729n.$$

**Proof.** Let  $G \in \mathcal{G}(n, 1/2)$  be a random graph, with vertex set  $V = [n]$ , say. For  $A \subset V$  let  $X_A$  be the indicator variable of the event

$$\{G[A] \text{ has all degrees odd}\},$$

and, for  $x, y \in V$ , let  $E_{xy}$  be the indicator variable of the event  $\{xy \in E(G)\}$ . We define

$$X_i = \sum_{A \in \mathcal{V}^{(i)}} X_A.$$

If  $|A|$  is odd, then  $X_A = 0$ . If  $|A|$  is even, then pick  $a \in A$  and let  $A' = A \setminus \{a\}$ . Since  $\sum_{x \in A} d_{G[A]}(x)$  is even, we have  $X_A = 1$  iff  $d_{G[A]}(x)$  is odd for all  $x \in A'$ . We condition on  $G[A']$ : for  $x \in A'$ , the parity of  $d_{G[A]}(x)$  depends on  $E_{ax}$ , and these events are independent, so  $X_A = 1$  with probability  $2^{-|A'|} = 2^{1-|A|}$ . Thus if  $i$  is even, then

$$\mathbb{E}X_i = 2^{1-i} \binom{n}{i}.$$

To prove the required upper bound, suppose that  $t \geq s(n) + \omega(n)$ . It is easily checked that  $s(n) > n/2$ , so  $\mathbb{E}X_{s(n)+k} \leq 2^{-k}\mathbb{E}X_{s(n)}$  for  $k > 0$ . Therefore

$$\begin{aligned} \mathbb{P}(f(G) \geq t) &= \mathbb{P}(X_i > 0 \text{ for some } i \geq t) \\ &\leq \sum_{i=t}^n \mathbb{P}(X_i > 0) \leq \sum_{i=t}^n \mathbb{E}(X_i) \\ &\leq \sum_{i=t}^n 2^{i-s(n)} \mathbb{E}X_{s(n)} \leq 2^{1-\omega(n)}. \end{aligned}$$

Therefore  $\mathbb{P}(f(G) > s(n) + \omega(n)) \rightarrow 0$  as  $n \rightarrow \infty$ .

For the lower bound, we calculate the variance of  $X_i$ , where  $i$  is even. Let  $A, B \in V^{(i)}$  be such that  $A \neq B$ , and pick  $a \in A \setminus B$  and  $b \in B \setminus A$ . Let  $C = A \cup B \setminus \{a, b\}$ . Conditioning on  $G[C]$ , we see that, for  $x \in A \setminus \{a\}$ , the event

$$\{d_{G[A]}(x) \text{ is odd}\}$$

depends on  $E_{ax}$ , while, for  $x \in B \setminus \{b\}$ , the event

$$\{d_{G[B]}(x) \text{ is odd}\}$$

depends on  $E_{bx}$ . These events are all independent, so

$$\mathbb{P}(X_A = X_B = 1) = 2^{-|C|} = 2^{2-2i} = \mathbb{P}(X_A = 1)\mathbb{P}(X_B = 1),$$

regardless of  $G[C]$ . Thus the events  $X_A$ , for  $A \in V^{(i)}$ , are pairwise independent. This means that the variance turns out to be surprisingly small. Indeed,

$$\begin{aligned} \text{var}(X_i) &= \sum_{A \in V^{(i)}} \sum_{B \in V^{(i)}} \mathbb{E}(X_A X_B) - \mathbb{E}(X_A)\mathbb{E}(X_B) \\ &= \sum_{A \in V^{(i)}} \mathbb{E}(X_A) - \mathbb{E}(X_A)^2 \\ &\leq \sum_{A \in V^{(i)}} \mathbb{E}(X_A) = \mathbb{E}(X_i). \end{aligned}$$

By Chebyshev's inequality, we have

$$\mathbb{P}(X_i = 0) \leq \text{var}(X_i)/\mathbb{E}(X_i)^2 \leq 1/\mathbb{E}(X_i).$$

Now, we may assume that  $s(n) - \omega(n)$  is even, and that  $\omega(n) \rightarrow \infty$  slowly enough so that  $s(n) - \omega(n) > n/2$ . Therefore

$$\begin{aligned}\mathbb{E}(X_{s(n)-\omega(n)}) &\geq 2^{\omega(n)-2} \mathbb{E}(X_{s(n)-2}) \\ &\geq 2^{\omega(n)-2},\end{aligned}$$

by the definition of  $s(n)$ . Thus  $\mathbb{P}(X_{s(n)-\omega(n)} = 0) \rightarrow 0$  as  $n \rightarrow \infty$ , so  $\mathbb{P}(f(G) < s(n) - \omega(n)) \rightarrow 0$  as  $n \rightarrow \infty$ .  $\blacksquare$

The same arguments, with a little modification, give essentially the same result for induced subgraphs with all degrees even (note that such a subgraph can have any order, in contrast to the odd case), thus yielding the following analogue of Corollary B for random graphs.

**Theorem 15.** *Let  $s(n)$  be the smallest positive integer such that  $2^{1-s} \binom{n}{s} \leq 1$ , and let  $\omega(n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Then almost every  $G \in \mathcal{G}(n, \frac{1}{2})$  satisfies the following.*

- (a) *There is a set  $S \subset V(G)$  such that  $|S| = \lceil s(n) - \omega(n) \rceil$  and  $G[S]$  has all degrees even.*
- (b) *There is no set  $S \subset V(G)$  such that  $|S| \geq s(n) + \omega(n)$  and  $G[S]$  has all degrees even.*

**Acknowledgments.** I would like to thank Dr Y. Caro for letting me have a copy of his paper [2], and Dr Jamie Radcliffe for his helpful comments.

## References

- [1] B. Bollobás, *Graph Theory, An Introductory Course*, Springer-Verlag, New York, Heidelberg, Berlin, 1979, x+180pp.
- [2] Y. Caro, On induced subgraphs with odd degrees. To appear.
- [3] Y. Caro, I. Krasikov and Y. Roditty, On induced subgraphs of trees with restricted degrees. *Discrete Mathematics*. To appear.
- [4] L. Lovász, *Combinatorial Problems and Exercises*, North-Holland, Amsterdam, 1979, 551pp.
- [5] A.J. Radcliffe and A.D. Scott, Every tree has a large induced subgraph with all degrees odd. To appear.