

# Induced subgraph density. V. All paths approach Erdős-Hajnal

Tung Nguyen<sup>1</sup>  
Princeton University,  
Princeton, NJ 08544, USA

Alex Scott<sup>2</sup>  
Oxford University,  
Oxford, UK

Paul Seymour<sup>3</sup>  
Princeton University,  
Princeton, NJ 08544, USA

July 4, 2022; revised January 23, 2024

<sup>1</sup>Supported by AFOSR grants A9550-19-1-0187 and FA9550-22-1-0234, and by NSF grants DMS-1800053 and DMS-2154169.

<sup>2</sup>Supported by EPSRC grant EP/X013642/1

<sup>3</sup>Supported by AFOSR grants A9550-19-1-0187 and FA9550-22-1-0234, and by NSF grants DMS-1800053 and DMS-2154169.

## Abstract

The Erdős-Hajnal conjecture says that for every graph  $H$ , there exists  $c > 0$  such that every  $H$ -free graph  $G$  has a clique or stable set of size at least  $|G|^c$ . (A graph is “ $H$ -free” if no induced subgraph is isomorphic to  $H$ ). In this paper we are concerned with the case when  $H$  is a path. The conjecture has been proved for paths with at most five vertices, but not for longer paths. We prove that the conjecture is “nearly” true for all paths: for every path  $H$ , every  $H$ -free graph  $G$  has a clique or stable set of size at least  $2^{(\log |G|)^{1-o(1)}}$ .

As in previous papers of this series, we can strengthen this, replacing the hypothesis that  $G$  is  $H$ -free with the weaker hypothesis that  $G$  does not contain “many” copies of  $H$ , and strengthening the conclusion, replacing the large clique or stable set outcome with a “near-polynomial” version of Nikiforov’s theorem.

# 1 Introduction

Some terminology and notation: if  $G$  is a graph,  $G[X]$  denotes the induced subgraph with vertex set  $X$  of a graph  $G$ ;  $|G|$  denotes the number of vertices of  $G$ ;  $\overline{G}$  is the complement graph of  $G$ ; and a graph is  $H$ -free if it has no induced subgraph isomorphic to  $H$ . A famous conjecture of Erdős and Hajnal [3, 4] from 1977 says:

**1.1 The Erdős-Hajnal Conjecture:** *For every graph  $H$  there exists  $c > 0$  such that every  $H$ -free graph  $G$  has a stable set or clique of size at least  $|G|^c$ .*

This remains open, and has been proved only for a very limited set of graphs  $H$  (although see [7, 8] for a variety of new graphs  $H$  that satisfy 1.1).

If  $H$  is a graph, for each  $n > 0$  let  $f_H(n)$  be the largest integer such that every  $H$ -free graph with at least  $n$  vertices has a stable set or clique with size at least  $f_H(n)$ . Thus, the Erdős-Hajnal conjecture says that

**1.2 Conjecture:** *For every graph  $H$  there exists  $c > 0$  such that  $f_H(n) \geq n^c$  for all  $n > 0$ .*

What can we actually prove about  $f_H(n)$ ? Erdős and Hajnal [4] proved:

**1.3** *For every graph  $H$ , there exists  $c > 0$  such that  $f_H(n) \geq 2^{c(\log n)^{1/2}}$  for all  $n > 0$ .*

In [2], with Bucić, we improved this, showing:

**1.4** *For every graph  $H$ , there exists  $c > 0$  such that  $f_H(n) \geq 2^{c(\log n \log \log n)^{1/2}}$  for all  $n > 0$ .*

For general graphs  $H$ , this is the best bound known.

But the case when  $H$  is a path is of particular interest. Paths are very simple graphs, and yet until recently, the Erdős-Hajnal conjecture was open when  $H$  is the five-vertex path  $P_5$ . This was the smallest graph for which the conjecture was open, and therefore was a focus of attention. In a substantial breakthrough, Blanco and Bucić [1] improved  $1/2$  to  $2/3$  for  $P_5$ :

**1.5** *There exists  $c > 0$  such that  $f_{P_5}(n) \geq 2^{c(\log n)^{2/3}}$  for all  $n > 0$ .*

Even more recently, in [8], we showed that  $P_5$  satisfies the Erdős-Hajnal conjecture, superceding 1.5; but the Erdős-Hajnal conjecture for general paths remains open, and it is natural to ask whether at least some analogue of 1.5 can be proved for general paths.

In the present paper, we do better than that: in 1.5, with  $P_5$  replaced by any path,  $2/3$  can be replaced by any  $d < 1$  ( $d = 1$  is the Erdős-Hajnal conjecture itself). More exactly:

**1.6** *For every path  $P$ , and all  $d < 1$ , there exists  $c > 0$  such that  $f_P(n) \geq 2^{c(\log n)^d}$  for all  $n > 0$ .*

This is equivalent to saying that for every path  $P$ ,  $f_P(n) \geq 2^{(\log n)^{1-o(1)}}$ .

This will be further strengthened, in two ways both of which need more definitions. If  $\varepsilon > 0$ , a subset  $S \subseteq V(G)$  is

- $\varepsilon$ -sparse if  $G[S]$  has maximum degree at most  $\varepsilon|S|$ ;
- $(1 - \varepsilon)$ -dense if  $\overline{G}[S]$  is  $\varepsilon$ -sparse, where  $\overline{G}$  is the complement graph of  $G$ ; and

- $\varepsilon$ -restricted if  $S$  is either  $\varepsilon$ -sparse or  $(1 - \varepsilon)$ -dense.

A (mildly strengthened) result of Rödl [10] says:

**1.7** For all  $0 < \varepsilon \leq 1/2$ , there exists  $\delta > 0$  such that for every  $H$ -free graph  $G$ , there is an  $\varepsilon$ -restricted subset  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ .

Fox and Sudakov [5] proposed the conjecture that the dependence of  $\delta$  on  $\varepsilon$  is polynomial; or more exactly:

**1.8 Conjecture:** For every graph  $H$  there exists  $c > 0$  such that for every  $\varepsilon$  with  $0 < \varepsilon \leq 1/2$  and every  $H$ -free graph  $G$ , there exists  $S \subseteq V(G)$  with  $|S| \geq \varepsilon^c|G|$  such that  $S$  is  $\varepsilon$ -restricted.

Every graph  $H$  satisfying this also satisfies the Erdős-Hajnal conjecture; and in the converse direction, it is proved in [6, 7, 8] that all the graphs currently known to satisfy the Erdős-Hajnal conjecture also satisfy conjecture 1.8.

As we said, 1.6 will be strengthened in two ways. The first strengthening is, we will prove that every path “nearly” satisfies the Fox-Sudakov conjecture. More exactly:

**1.9** For every path  $P$  and all  $d < 1$ , there exists  $c > 0$  such that for all  $0 < \varepsilon \leq 1/2$  and every  $P$ -free graph  $G$ , there is an  $\varepsilon$ -restricted subset  $S \subseteq V(G)$  with  $|S| \geq 2^{-c(\log \frac{1}{\varepsilon})^{1/d}}|G|$ .

This first strengthening is crucial to the proof of 1.6.

The second strengthening is, we can replace the hypothesis of 1.9 that  $G$  is  $P$ -free, with a weaker hypothesis that  $G$  does not contain many copies of  $P$ . A *copy* of  $H$  in  $G$  is an isomorphism from  $H$  to an induced subgraph of  $G$ . Let  $\text{ind}_H G$  be the number of copies of  $H$  in  $G$ . There is a theorem of Nikiforov [9], strengthening Rödl’s theorem:

**1.10** For every graph  $H$  and all  $\varepsilon > 0$ , there exists  $\delta > 0$  such that for every graph  $G$ , if  $\text{ind}_H(G) \leq (\delta|G|)^{|H|}$ , then there is an  $\varepsilon$ -restricted subset  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ .

We will prove that, when  $H$  is a path, this is satisfied taking  $\delta$  to be a “near-polynomial” function of  $\varepsilon$ . More exactly:

**1.11** For every path  $P$  and all  $d < 1$ , there exists  $c > 0$  such that for all  $0 < \varepsilon \leq 1/2$ , if  $\delta$  satisfies

$$\delta = 2^{-c(\log \frac{1}{\varepsilon})^{1/d}},$$

then in every graph  $G$  with  $\text{ind}_P(G) \leq (\delta|G|)^{|P|}$ , there is an  $\varepsilon$ -restricted subset  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ .

This second strengthening is not crucial for the proof, but the proof method gives it naturally anyway. It also has the advantage that the class of graphs  $P$  (not necessarily paths) that satisfy 1.11 is closed under vertex-substitution, while we cannot prove the same for the the class of graphs  $P$  that satisfy 1.9.

Thus, 1.11 is our main result. Obviously it implies 1.9, but that it implies 1.6 is not so obvious. Let us see that, via the following.

**1.12** Let  $G$  be a graph with  $|G| \geq 2$ , and assume that  $a > 0$  is such that for every  $\varepsilon > 0$  with  $\varepsilon \leq 1/2$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with

$$|S| \geq 2^{-a(\log \frac{1}{\varepsilon})^{1/d}} |G|.$$

Then  $G$  contains a clique or stable set of size at least  $2^{c(\log |G|)^d}$ , where  $c = (2a + 2)^{-1}$ .

**Proof.** If  $2^{c(\log |G|)^d} \leq 2$ , then the result holds since  $G$  has a clique or stable set of size two; so we assume that  $2^{c(\log |G|)^d} > 2$ . Let  $\varepsilon := 2^{-2c(\log |G|)^d}$ ; then  $\varepsilon \in (0, \frac{1}{4})$ . Let  $\delta > 0$  be such that

$$\log \frac{1}{\delta} = a \left( \log \frac{1}{\varepsilon} \right)^{1/d} = a(2c)^{1/d} \log |G| \leq 2ac \log |G|.$$

Then  $\delta \geq |G|^{-2ac}$ . From the hypothesis, there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with

$$|S| \geq \delta |G| \geq |G|^{1-2ac} = |G|^{2c} = 2^{2c \log |G|} \geq 2^{2c(\log |G|)^d} = \varepsilon^{-1}.$$

Thus, since  $S$  is  $\varepsilon$ -restricted,  $G[S]$  (and so  $G$ ) contains a clique or stable set of size at least

$$\frac{|S|}{\varepsilon |S| + 1} \geq \frac{1}{2\varepsilon} \geq \varepsilon^{-1/2} = 2^{c(\log |G|)^d}.$$

This proves 1.12. ■

## 2 Blockades

As in previous papers of this series, we say a graph  $H$  is *viral* if there exists  $c > 0$  such that for all  $0 < \varepsilon \leq 1/2$  and every graph  $G$  with  $\text{ind}_H(G) \leq (\varepsilon^c |G|)^{|H|}$ , there is an  $\varepsilon$ -restricted subset  $S \subseteq V(G)$  with  $|S| \geq \varepsilon^c |G|$ . Let us say that a graph  $H$  is *near-viral* if for every  $d < 1$ , there exists  $c > 0$  such that for every  $\varepsilon \in (0, \frac{1}{2})$ , if  $\delta$  satisfies

$$\log \frac{1}{\delta} = c \left( \log \frac{1}{\varepsilon} \right)^{1/d},$$

then for every graph  $G$  with  $\text{ind}_H(G) \leq (\delta |G|)^{|H|}$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta |G|$ . Thus, our main theorem 1.11 says that every path is near-viral.

If  $A, B \subseteq V(G)$  are disjoint, we say  $A$  is  $x$ -sparse to  $B$  if every vertex in  $A$  has at most  $\varepsilon |B|$  neighbours in  $B$ , and  $A$  is  $(1-x)$ -dense to  $B$  if every vertex in  $A$  has at least  $(1-\varepsilon) |B|$  neighbours in  $B$ . A *blockade* in a graph  $G$  is a finite sequence  $(B_1, \dots, B_n)$  of (possibly empty) disjoint subsets of  $V(G)$ ; its *length* is  $n$  and its *width* is  $\min_{i \in [n]} |B_i|$ . For  $k, w \geq 0$ ,  $(B_1, \dots, B_n)$  is a  $(k, w)$ -blockade if its length is at least  $k$  and its width is at least  $w$ . For  $x \in (0, \frac{1}{2})$ , this blockade is  $x$ -sparse if  $B_j$  is  $x$ -sparse to  $B_i$  for all  $i, j \in [n]$  with  $i < j$ , and  $(1-x)$ -dense if  $B_j$  is  $(1-x)$ -dense to  $B_i$  for all  $i, j \in [n]$  with  $i < j$ .

Thus, there are three parameters we care about, the length, width, and sparsity (or density). It is easier to prove that certain graphs contain blockades with some desired combination of the three

parameters, than to prove directly that they contain large  $\varepsilon$ -restricted sets. But the reason blockades are useful for us is 2.3 below, that says that if a graph  $G$  and all its large induced subgraphs admit blockades with certain parameters, then  $G$  must contain a large  $\varepsilon$ -restricted set.

There are now several theorems of this type, with a family resemblance, but sufficiently different to be confusing, and perhaps it would be helpful to summarize them here.

- Erdős and Hajnal [4] proved that for every graph  $H$ , there exists  $d > 0$  such that for all  $x \in (0, 1/2]$ , every  $H$ -free graph admits an  $x$ -sparse or  $(1-x)$ -dense  $(2, \lfloor x^d |G| \rfloor)$  blockade. From this they deduced their result 1.3.
- In [2], we (with Bucić) proved a strengthening, that for every graph  $H$ , there exists  $d > 0$  such that for all  $x \in (0, 1/2]$ , every  $H$ -free graph (or every graph  $G$  with  $\text{ind}_H(G) \leq (x^d |G|)^{|H|}$ ) admits an  $x$ -sparse or  $(1-x)$ -dense  $(\log(1/x), \lfloor x^d |G| \rfloor)$ -blockade. This allowed us to deduce 1.4.
- If we could prove that for a graph  $H$ , there exists  $d > 0$  such that for all  $x \in (0, 1/2]$ , every graph  $G$  with  $\text{ind}_H(G) \leq (x^d |G|)^{|H|}$  admits an  $x$ -sparse or  $(1-x)$ -dense  $(1/x, \lfloor x^d |G| \rfloor)$ -blockade, then we could deduce that  $H$  is viral. It would be just as good if we could prove that for all  $x \in (0, 1/2]$ , every graph  $G$  with  $\text{ind}_H(G) \leq (x^d |G|)^{|H|}$  admits an  $x$ -sparse or  $(1-x)$ -dense  $(k, \lfloor |G|/k^d \rfloor)$ -blockade for some  $k \in [2, 1/x]$ . This is used in [6] to prove that  $C_5$  is viral.
- Suppose that there exist  $a, b > 0$  and  $d > 2$  such that for all  $0 < x < y \leq 1/2$ , every  $y^a$ -restricted graph  $G$  with  $\text{ind}_H(G) \leq (x^{bd^2} |G|)^{|H|}$  admits either a  $y^{ad}$ -restricted subset of size at least  $y^{bd^2} |G|$ , or an  $x$ -sparse or  $(1-x)$ -dense  $(1/y, \lfloor y^{bd^2} |G| \rfloor)$ -blockade. Then  $H$  is viral. This was the method used in [7].
- Suppose we could prove that there exists  $d > 0$  such that for all  $x, y$  with  $0 < x < y \leq 1/2$ , every poly( $y$ )-restricted graph  $G$  with  $\text{ind}_H(G) \leq (x^d |G|)^{|H|}$  admits an  $x$ -sparse or  $(1-x)$ -dense  $(1/y, \lfloor x^d |G| \rfloor)$ -blockade. Then we could deduce that  $H$  is near-viral. This is the approach in this paper.

The *edge-density* of a graph  $G$  is  $|E(G)|$  divided by  $\binom{|G|}{2}$  (or 1 if  $|G| \leq 1$ ). Let us say a subset  $S \subseteq V(G)$  is *weakly  $\varepsilon$ -restricted* if one of  $G[S], \overline{G}[S]$  has edge-density at most  $\varepsilon$ . A function  $\ell: (0, \frac{1}{2}) \rightarrow \mathbb{R}^+$  is *subreciprocal* if it is nonincreasing and  $1 < \ell(x) \leq 1/x$  for all  $x \in (0, \frac{1}{2})$ . For a subreciprocal function  $\ell$ , a graph  $H$  is  *$\ell$ -divisive* if there are  $c \in (0, \frac{1}{2})$  and  $d > 1$  such that for every  $x \in (0, c)$  and every graph  $G$  with  $\text{ind}_H(G) \leq (x^d |G|)^{|H|}$ , there is an  $x$ -sparse or  $(1-x)$ -dense  $(\ell(x), \lfloor x^d |G| \rfloor)$ -blockade in  $G$ . The next result is proved in [2].

**2.1** *Let  $\ell: (0, \frac{1}{2}) \rightarrow \mathbb{R}^+$  be subreciprocal, and let  $H$  be  $\ell$ -divisive. Then there exists  $C > 0$  such that for every  $\varepsilon \in (0, \frac{1}{2})$ , if we define  $\delta > 0$  by*

$$\log \frac{1}{\delta} = \frac{C(\log \frac{1}{\varepsilon})^2}{\log(\ell(\varepsilon))},$$

*then for every graph  $G$  with  $\text{ind}_H(G) \leq (\delta |G|)^{|H|}$ , there is a weakly  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta |G|$ .*

This provides us with large subsets that are weakly  $\varepsilon$ -restricted, but we want  $\varepsilon$ -restricted subsets. These are easy to find, because of the following lemma:

**2.2** For  $\varepsilon \in (0, \frac{1}{2})$  and a graph  $G$ , if  $S \subseteq V(G)$  is weakly  $\frac{1}{4}\varepsilon$ -restricted, then there exists an  $\varepsilon$ -restricted  $T \subseteq S$  with  $|T| \geq \frac{1}{2}|S|$ .

**Proof.** We may assume that  $G[S]$  has at most  $\frac{1}{4}\varepsilon \binom{|S|}{2} < \frac{1}{8}\varepsilon|S|^2$  edges. Let  $T$  be the set of vertices in  $S$  with degree at most  $\frac{1}{2}\varepsilon|S|$  in  $G[S]$ ; then  $\frac{1}{2}\varepsilon|S||S \setminus T| < \frac{1}{4}\varepsilon|S|^2$  and so  $|S \setminus T| < \frac{1}{2}|S|$ . Thus  $|T| > \frac{1}{2}|S|$  and  $G[T]$  has maximum degree at most  $\frac{1}{2}\varepsilon|S| < \varepsilon|T|$ . This proves 2.2. ■

So we obtain a version of 2.1 that gives  $\varepsilon$ -restricted sets:

**2.3** Let  $\ell$  be subreciprocal, and let  $H$  be an  $\ell$ -divisive graph. Then there exists  $C > 0$  such that for every  $\varepsilon \in (0, \frac{1}{2})$ , if we define  $\delta > 0$  by

$$\log \frac{1}{\delta} = \frac{C(\log \frac{1}{\varepsilon})^2}{\log(\ell(\varepsilon))},$$

then for every graph  $G$  with  $\text{ind}_H(G) \leq (\delta|G|)^{|H|}$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ .

**Proof.** Choose  $C'$  such that 2.1 holds with  $C$  replaced by  $C'$ . We claim that  $C := 18C'$  satisfies the theorem. To show this, let  $\varepsilon \in (0, \frac{1}{2})$ , let  $\delta$  be as in 2.3, and let  $G$  be a graph with  $\text{ind}_H(G) \leq (\delta|G|)^{|H|}$ . We must show that there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ .

Let  $\varepsilon' := \frac{1}{4}\varepsilon \in (\varepsilon^3, \varepsilon)$ , and define  $\delta'$  by

$$\log \frac{1}{\delta'} := \frac{C'(\log \frac{1}{\varepsilon'})^2}{\log(\ell(\varepsilon'))} < \frac{C'(\log \frac{1}{\varepsilon^3})^2}{\log(\ell(\varepsilon'))} \leq \frac{9C'(\log \frac{1}{\varepsilon})^2}{\log(\ell(\varepsilon))} = \frac{1}{2} \log \frac{1}{\delta},$$

and so  $\delta' \geq \sqrt{\delta} > 2\delta$ . Since  $\text{ind}_H(G) \leq (\delta|G|)^{|H|} \leq (\delta'|G|)^{|H|}$ , there is a weakly  $\varepsilon'$ -restricted  $S \subseteq V(G)$  in  $G$  with  $|S| \geq \delta'|G|$ . By 2.2, there is an  $\varepsilon$ -restricted  $T \subseteq S$  with  $|T| \geq \frac{1}{2}|S| \geq \frac{1}{2}\delta'|G| > \delta|G|$ . This proves 2.3. ■

For each integer  $s \geq 0$ , let  $\ell_s: (0, \frac{1}{2}) \rightarrow \mathbb{R}^+$  be the function defined by

$$\ell_s(x) := 2^{(\log \frac{1}{x})^{\frac{s}{s+1}}}$$

for all  $x \in (0, \frac{1}{2})$ . Then  $\ell_s$  is subreciprocal. We will show that:

**2.4** Every path  $P$  is  $\ell_s$ -divisive for all integers  $s \geq 0$ .

Let us deduce 1.11, which we restate:

**2.5** Every path  $P$  is near-viral.

**Proof (assuming 2.4).** We must show that for all  $d < 1$ , there exists  $c > 0$  such that for all  $\varepsilon \in (0, 1/2)$ , if  $\delta$  satisfies

$$\log \frac{1}{\delta} = c \left( \log \frac{1}{\varepsilon} \right)^{1/d},$$

then for every graph  $G$  with  $\text{ind}_H(G) \leq (\delta|G|)^{|H|}$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ .

Choose  $s$  with  $\frac{s+1}{s+2} \geq d$ . Since  $P$  is  $\ell_s$ -divisive, by 2.3 there exists  $C > 0$  such that for every  $\varepsilon \in (0, \frac{1}{2})$ , if we define  $\delta' > 0$  by

$$\log \frac{1}{\delta'} = \frac{C \left(\log \frac{1}{\varepsilon}\right)^2}{\log(\ell_s(\varepsilon))} = \frac{C \left(\log \frac{1}{\varepsilon}\right)^2}{\left(\log \left(\frac{1}{\varepsilon}\right)\right)^{s/(s+1)}} = C \left(\log \frac{1}{\varepsilon}\right)^{\frac{s+2}{s+1}},$$

then for every graph  $G$  with  $\text{ind}_H(G) \leq (\delta'|G|)^{|H|}$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta'|G|$ . We claim that we make take  $c = C$ . To see this, let  $\varepsilon \in (0, 1/2)$ , let  $\delta$  satisfy

$$\log \frac{1}{\delta} = C \left(\log \frac{1}{\varepsilon}\right)^{1/d},$$

and let  $G$  be a graph with  $\text{ind}_H(G) \leq (\delta|G|)^{|H|}$ . Then

$$\log \frac{1}{\delta'} = C \left(\log \frac{1}{\varepsilon}\right)^{\frac{s+2}{s+1}} \leq C \left(\log \frac{1}{\varepsilon}\right)^{1/d} = \log \frac{1}{\delta}$$

and so  $\delta' \geq \delta$ . Since  $\text{ind}_P(G) \leq (\delta'|G|)^k$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta'|G| \geq \delta|G|$ . This proves 2.5.  $\blacksquare$

### 3 In a sparse graph

The remainder of the paper is devoted to the proof of 2.4. Its proof proceeds by induction on  $s$ ; so we may assume that the path  $P$  is  $\ell_{s-1}$ -divisive. We are given a graph  $G$  with  $\text{ind}_P(G)$  small, and we need to find a blockade in  $G$  with certain parameters, to show that  $P$  is  $\ell_s$ -divisive. It follows from 2.3 that  $V(G)$ , and every large subset of  $V(G)$ , includes a somewhat smaller subset  $S$  that is either very dense or very sparse.

In this section we prove a result that allows us to win if ever the subset  $S$  turns out to be sparse. There is no symmetry between the dense and sparse cases, because we are excluding a path but not its complement; and as usual with problems about excluding a path, our task is easier if the host graph is sparse, when we can use a modified version of the well-known ‘‘Gyarfas path argument’’. We will prove:

**3.1** *Let  $P$  be a path, let  $0 < x \leq y \leq 1/(2|P|)$ , and let  $G$  be a  $y^2$ -sparse graph. Then either:*

- $\text{ind}_P(G) \geq (x^4|G|)^{|P|}$ , or
- *there is an  $x$ -sparse  $(1/y, \lfloor x^5|G| \rfloor)$ -blockade in  $G$ .*

**Proof.** Let  $|P| = k \geq 1$ . If  $k = 1$  the first bullet holds, and if  $x^5|G| < 1$  then the second bullet holds, so we assume that  $k \geq 2$  and  $|G| \geq x^{-5}$ . Choose an  $x$ -sparse blockade  $(B_1, \dots, B_{n-1}, C)$  in  $G$  with  $n$  maximum such that  $|B_1|, \dots, |B_{n-1}|, |C| \geq x^5|G|$  and  $|C| \geq (1 - k(n-1)y^2)|G|$ . We may assume that  $n < 1/y$ , and so

$$|C| \geq (1 - k(n-1)y^2)|G| = (1 + ky^2 - kny^2)|G| \geq (1/2 + ky^2)|G|.$$



We claim:

(1) For every  $X \subseteq C$  with  $|X| \geq x^4|G|$ , and  $Y \subseteq C \setminus X$  with  $|Y| \geq (1 + 4x^3 - kny^2)|G|$ , some vertex in  $X$  has at least  $2x^4|G|$  neighbours in  $Y$ .

Suppose not. Then  $|Y| \geq (1 + 4x^3 - kny^2)|G| \geq |G|/2$ , since  $kny^2 \leq 1/2$ . There are most  $2x^4|G| \cdot |X| \leq 4x^4|X| \cdot |Y|$  edges between  $X$  and  $Y$ , and so at most  $4x^3|G|$  vertices in  $Y$  have at least  $x|X|$  neighbours in  $X$ . Thus there is a subset  $Y'$  of  $Y$  with cardinality at least

$$|Y| - 4x^3|G| \geq (1 - kny^2)|G| \geq |G|/2 \geq x^5|G|$$

that is  $x$ -sparse to  $X$ . But then  $(B_1, \dots, B_{n-1}, X, Y')$  contradicts the maximality of  $n$ . This proves (1).

For  $t \geq 1$  an integer, let us say a  $t$ -brush is an induced path  $v_1 \cdots v_t$  of  $G[C]$ , such that  $v_t$  has at least  $2x^4|G|$  neighbours in  $C$  that are different from and nonadjacent to each of  $v_1, \dots, v_{t-1}$ .

(2) For  $1 \leq t \leq k - 2$ , if  $v_1 \cdots v_t$  is a  $t$ -brush of  $G[C]$ , there are at least  $x^4|G|$  vertices  $v$  such that  $v_1 \cdots v_t v$  is a  $(t + 1)$ -brush.

Let  $X$  be the set of neighbours of  $v_t$  in  $C$  that are different from and nonadjacent to each of  $v_1, \dots, v_{t-1}$ ; and let  $Y$  be the set of all vertices in  $C$  that are different from and nonadjacent to each of  $v_1, \dots, v_t$ . Thus  $|X| \geq 2x^4|G|$  since  $v_1 \cdots v_t$  is a  $t$ -brush. Moreover,  $k \geq 3$  since  $1 \leq t \leq k - 2$ , and so, as  $G$  is  $y^2$ -sparse,

$$|Y| \geq |C| - (k - 2)y^2|G| \geq (1 - k(n - 1)y^2 - (k - 2)y^2)|G| = (1 + 2y^2 - kny^2)|G| \geq (1 + 4x^3 - kny^2)|G|$$

(because  $4x^3 \leq 2y^2$ ). By (1), fewer than  $x^4|G|$  vertices in  $X$  have fewer than  $2x^4|G|$  neighbours in  $Y$ . All the others give  $(t + 1)$ -brushes extending  $v_1 \cdots v_t$ ; and since  $|X| - x^4|G| \geq x^4|G|$ , this proves (2).

(3) There are at least  $|G|/2$  1-brushes.

Suppose there is a set  $X$  of  $\lceil x^4|G| \rceil$  vertices in  $C$  each with degree less than  $2x^4|G|$  in  $G[C]$ . Let  $Y = C \setminus X$ ; then

$$|Y| \geq |C| - x^4|G| - 1 \geq |C| - x^4|G| - x^5|G| \geq (1 - k(n - 1)y^2 - x^4 - x^5)|G| \geq (1 + 4x^3 - kny^2)|G|$$

(since  $kny^2 \geq 2x^2 \geq 4x^3 + x^4 + x^5$ ), contrary to (1). Thus there are fewer than  $x^4|G|$  vertices that have degree less than  $2x^4|G|$  in  $G[C]$ . All the others give 1-brushes, and since  $|C| - x^4|G| \geq |G|/2$ , this proves (3).

From (2) and (3), it follows inductively that for  $1 \leq t \leq k - 1$  there are at least  $x^{4(t-1)}|G|^{t/2}/2$   $t$ -brushes, and in particular, there are at least  $x^{4(k-1)}|G|^{k-1}/2$   $(k - 1)$ -brushes. Each extends to at least  $2x^4|G|$  induced  $k$ -vertex paths; and so  $\text{ind}_P(G) \geq (2x^4|G|)x^{4(k-1)}|G|^{k-1}/2 = x^{4k}|G|^k$ . This proves 3.1. ■

## 4 The dense case

As we discussed at the start of the previous section, for the inductive proof of 2.4, we will now be able to assume that every large subset of  $V(G)$  includes a somewhat smaller subset that is very dense. That motivates the following:

**4.1** *Let  $P$  be a path with  $|P| \geq 1$ , and let  $0 < x \leq y \leq 1/100$ . Let  $G$  be a graph such that for every  $S \subseteq V(G)$  with  $|S| \geq x^{3|P|}|G|$ , there is a  $(1 - y^3)$ -dense subset  $S' \subseteq S$  with  $|S'| \geq x|S|$ . Then either:*

- $\text{ind}_P(G) \geq (x^{3|P|}|G|)^{|P|}$ ; or
- *there is a  $(1 - x)$ -dense  $(1/y, \lfloor x^{3|P|}|G| \rfloor)$ -blockade in  $G$ .*

**Proof.** In the proof of 3.1, we counted “ $t$ -brushes”, induced  $t$ -vertex paths in which the last vertex had many neighbours that all had no neighbours in the earlier part of the path. The issue there was to prove that, given a  $t$ -brush, there were many ways to extend it to a  $(t + 1)$ -brush. We will do something similar here, but we need to redefine a  $t$ -brush. We will be working inside a graph that is very dense, so there is no problem arranging that the last vertex of the path has many neighbours; the issue is to arrange that there are many vertices with no neighbours in the path, and to maintain this as we grow the path. A *non-neighbour* of  $v$  means a vertex different from and nonadjacent to  $v$ , and the *antidegree* of  $v$  is the number of its non-neighbours.

Let  $k := |P|$  and  $a := 3k$ . We may assume that  $x^a|G| \geq 1$ , since otherwise the second bullet holds. Define  $a_1 := x/2$ , and  $b_1 := x^2y/8$ ; and for  $2 \leq t \leq k$ , define  $a_t := (x/2)b_{t-1}$  and  $b_t := (x^2/2)b_{t-1}$ . For  $1 \leq t \leq k$  let us say a  *$t$ -brush* is an induced path of  $G$  with vertices  $v_1 \cdots v_t$  in order, such that there exist subsets  $A, B \subseteq V(G)$  with the following properties:

- every vertex in  $A$  is adjacent to  $v_t$  and is nonadjacent to  $v_1, \dots, v_{t-1}$ ;
- every vertex in  $B$  has no neighbours in  $\{v_1, \dots, v_t\}$ ;
- $|A| \geq a_t|G|$  and  $|B| \geq b_t|G|$ ;
- for every  $Y \subseteq B$  with  $|Y| \geq x^a|G|$ , there are at least  $y|A|/4$  vertices in  $A$  that have at least  $x|Y|$  non-neighbours in  $Y$ ; and
- every vertex in  $B$  has at most  $3y^3|A|$  non-neighbours in  $A$ .

We claim first:

(1) *For every  $S \subseteq V(G)$  with  $|S| \geq 2x^{a-1}|G|$ , there exists  $C \subseteq S$  with  $|C| \geq x(1 + y)|S|/2$ , such that  $C$  is  $(1 - 2y^3)$ -dense, and for all disjoint  $X, Y \subseteq C$  with  $|X| \geq (1 - y/4)|C|$  and  $|Y| \geq x^a|G|$ , at least  $y|X|/4$  vertices in  $X$  have at least  $x|Y|$  non-neighbours in  $Y$ .*

Since  $|S| \geq x^a|G|$ , there exists  $S' \subseteq S$  with  $|S'| \geq x|S|$  such that  $S'$  is  $(1 - y^3)$ -dense. Choose a  $(1 - x)$ -dense blockade  $(B_1, \dots, B_{n-1}, C)$  in  $G[S']$  with  $n$  maximum such that  $|B_1|, \dots, |B_{n-1}|, |C| \geq x^a|G|$  and  $|C| \geq (1 - (n - 1)y/2)|S'|$ . (This is possible because  $|S'| \geq x^a|G|$ , and so we can take  $n = 1$  and  $C = S'$ .) We may assume that  $n < 1/y$ , and so

$$|C| \geq (1 - (n - 1)y/2)|S'| = (1 + y/2 - ny/2)|S'| \geq (1 + y)|S'|/2 \geq x(1 + y)|S|/2.$$

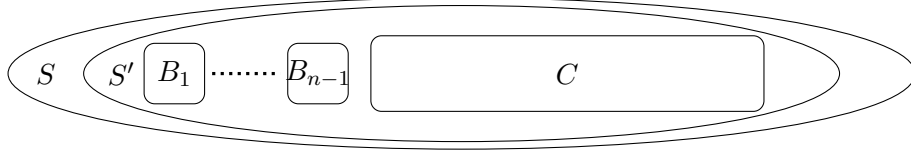


Figure 1: For step (1) of the proof of 4.1.

In particular,  $|C| \geq |S'|/2$ , and consequently  $C$  is  $(1 - 2y^3)$ -dense.

Suppose that  $X, Y \subseteq C$  are disjoint, with  $|X| \geq (1 - y/4)|C|$  and  $|Y| \geq x^a|G|$ . It follows that

$$|X| \geq (1 - y/4)(1 + y)|S'|/2 \geq |S'|/2.$$

Since  $|Y| \geq x^a|G|$ , and  $(1 - ny/2)|S'| \geq |S'|/2 \geq x^a|G|$ , fewer than  $(1 - ny/2)|S'|$  vertices in  $X$  are  $(1 - x)$ -dense to  $Y$ , from the maximality of  $n$ . Since  $|C| \geq (1 - (n - 1)y/2)|S'|$ , it follows that at least  $y|S'|/2 - |Y|$  vertices in  $X$  have at least  $x|Y|$  non-neighbours in  $Y$ . But  $|Y| \leq y|C|/4$ , since  $X \cap Y = \emptyset$ , and so  $y|S'|/2 - |Y| \geq y|S'|/2 - y|C|/4 \geq y|C|/4 \geq y|X|/4$ . This proves (1).

(2) *There are at least  $x|G|/2$  1-brushes.*

Since  $|G| \geq 2x^{a-1}|G|$ , (1) implies that there exists  $C \subseteq V(G)$  with  $|C| \geq x(1 + y)|G|/2$ , such that  $C$  is  $(1 - 2y^3)$ -dense, and for all disjoint  $X, Y \subseteq C$  with  $|X| \geq (1 - y/4)|C|$  and  $|Y| \geq x^a|G|$ , at least  $y|X|/4$  vertices in  $X$  have at least  $x|Y|$  non-neighbours in  $Y$ .

Suppose that there is a set  $Y$  of  $\lceil x^a|G| \rceil$  vertices in  $C$  each with antidegree less than  $(x^2y/8)|G|$  in  $G[C]$ . Let  $X = C \setminus Y$ . Then

$$|Y| \leq x^a|G| + 1 \leq 2x^a|G| \leq 4x^{a-1}|C| \leq (x/4)|C| \leq (y/4)|C|.$$

There are at most  $(x^2y/8)|G| \cdot |Y|$  nonedges between  $X, Y$ ; and yet from the choice of  $C$ , since  $|X| \geq (1 - y/4)|C|$ , there are at least  $(y|X|/4)(x|Y|) = (xy/4)|X| \cdot |Y|$  such nonedges. So

$$(xy/4)|X| \cdot |Y| \leq (x^2y/8)|G| \cdot |Y|,$$

and so  $2|X| \leq x|G|$ . But

$$|X| \geq |C| - x^a|G| - 1 \geq |C| - 2x^a|G| \geq x(1 + y)|G|/2 - 2x^a|G| > x|G|/2,$$

a contradiction.

Thus there are fewer than  $x^a|G|$  vertices that have antidegree less than  $(x^2y/8)|G|$  in  $G[C]$ ; and so there are at least  $|C| - x^a|G| \geq x|G|/2$  vertices in  $C$  with antidegree at least  $(x^2y/8)|G|$  in  $G[C]$ . We claim that each such vertex forms a 1-brush. Let  $v$  be such a vertex, and let  $A, B$  be its sets of neighbours and non-neighbours in  $G[C]$ . Then  $b_1|G| = (x^2y/8)|G| \leq |B| \leq 2y^3|C|$ , and so (since  $1 \leq x^a|G| \leq 2x^{a+1}|C| \leq y^3|C|$ )

$$|A| \geq |C| - 2y^3|C| - 1 \geq (1 - 3y^3)|C| \geq (1 - 3y^3)x(1 + y)|G|/2 \geq x|G|/2 = a_1|G|.$$

Moreover, since  $C$  is  $(1 - 2y^3)$ -dense, every vertex in  $B$  has at most  $2y^3|C| \leq 3y^3|A|$  non-neighbours in  $A$ . Finally, let  $Y \subseteq B$  with  $|Y| \geq x^a|G|$ . Since  $|A| \geq (1 - 3y^3)|C| \geq (1 - y/4)|C|$ , the choice of  $C$

implies that at least  $y|A|/4$  vertices in  $A$  have at least  $x|Y|$  non-neighbours in  $Y$ . Hence  $v$  forms a 1-brush. This proves (2).

(3) Let  $1 \leq t \leq k - 1$ , and let  $v_1 - \dots - v_t$  be a  $t$ -brush. Then there are at least  $ya_t|G|/8$  vertices  $v$  such that  $v_1 - \dots - v_t - v$  is a  $(t + 1)$ -brush.

Choose  $A, B$  satisfying the five bullets in the definition of “ $t$ -brush”. Since  $b_t = (x^2/2)^t y/4$ , and  $t \leq k - 1$ , and  $a \geq 3k$ , it follows that

$$|B| \geq b_t|G| = (x^2/2)^t y|G|/4 \geq x^{3k-4}|G| \geq 2x^{a-1}|G|.$$

By (1), there exists  $C \subseteq B$  with  $|C| \geq x(1 + y)|B|/2$ , such that  $C$  is  $(1 - 2y^3)$ -dense, and for all disjoint  $X, Y \subseteq C$  with  $|X| \geq (1 - y/4)|C|$  and  $|Y| \geq x^a|G|$ , at least  $y|X|/4$  vertices in  $X$  have at least  $x|Y|$  non-neighbours in  $Y$ .

Since  $v_1 - \dots - v_t$  is a  $t$ -brush, each vertex in  $C$  has at most  $3y^3|A|$  non-neighbours in  $A$ , and so at most  $y|A|/8$  vertices in  $A$  have at least  $24y^2|C|$  non-neighbours in  $C$ . On the other hand, there are at least  $y|A|/4$  vertices in  $A$  that have at least  $x|C|$  non-neighbours in  $C$ ; and so there is a set  $D \subseteq A$  with  $|D| \geq y|A|/8$ , such that for each  $v \in D$ , the number of its non-neighbours in  $C$  is between  $x|C|$  and  $24y^2|C|$ .

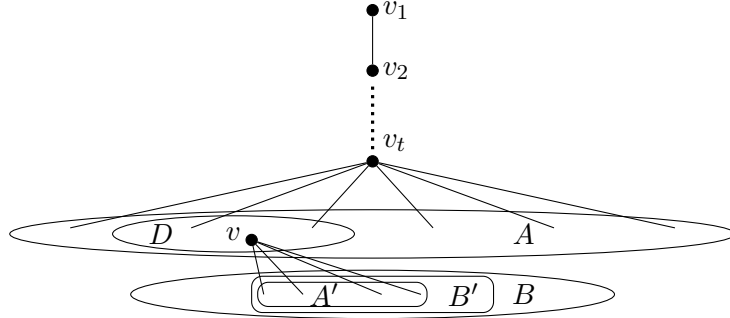


Figure 2: For step (3).  $C = A' \cup B'$ .

Let  $v \in D$ . We claim that  $v_1 - \dots - v_t - v$  is a  $(t + 1)$ -brush. Let  $A'$  be the set of all neighbours of  $v$  in  $C$ , and let  $B' = C \setminus A'$ . We will show that  $A', B'$  satisfy the five conditions in the definition of a  $(t + 1)$ -brush. The first two are immediate. For the third,

$$|A'| \geq (1 - 24y^2)|C| \geq (1 - 24y^2)x(1 + y)|B|/2 \geq (x/2)b_t|G| = a_{t+1}|G|,$$

and

$$|B'| \geq x|S| \geq x(x/2)|B| \geq (x^2/2)b_t|G| = b_{t+1}|G|.$$

For the fourth condition, suppose that  $Y \subseteq B'$  with  $|Y| \geq x^a|G|$ . From the choice of  $C$ , since  $|A'| \geq (1 - 24y^2)|C| \geq (1 - y/4)|C|$ , there are at least  $y|A'|/4$  vertices in  $A'$  that have at least  $x|Y|$  non-neighbours in  $Y$ . Finally, for the fifth condition, since  $C$  is  $(1 - 2y^3)$ -dense, each vertex in  $B'$  has at most  $2y^3|C| \leq 3y^3|A'|$  non-neighbours in  $A'$ . This proves (3).

(4) For  $1 \leq t \leq k$ , there are at least  $x^{3t^2}|G|^t$   $t$ -brushes.

This is true if  $t = 1$  by (2), so we assume inductively that  $2 \leq t \leq k$  and the result holds for  $t - 1$ , and we prove it for  $t$ . By (3), since  $t \geq 2$  and  $y \geq x$ , the number of  $t$ -brushes is at least  $xa_{t-1}|G|/8$  times the number of  $(t - 1)$ -brushes, and so at least  $a_{t-1}tx^{3(t-1)^2+1}|G|^t/8$ . If  $t = 2$ , the claim follows, so we assume that  $t \geq 3$ . From the definition,  $b_{t-2} = (x^2/2)^{t-3}(x^2y/8) \geq x^{2t-3}/2^t$ , and so

$$a_{t-1} = (x/2)b_{t-2} \geq x^{2t-2}/2^{t+1}.$$

Hence the number of  $t$ -brushes is at least

$$(x^{2t-2}/2^{t+1})x^{3(t-1)^2+1}|G|^t/8 = \left(x^{3t^2-4t+2}/2^{t+4}\right)|G|^t \geq x^{3t^2}|G|^t.$$

This proves (4).

From (4) with  $t = k$ , it follows that  $\text{ind}_{P_k}(G) \geq (x^{3k}|G|)^k$ . This proves 4.1.  $\blacksquare$

## 5 Decreasing density

We remind the reader that if  $\ell$  is subreciprocal, a graph  $H$  is  $\ell$ -divisive if there are  $c \in (0, \frac{1}{2})$  and  $d > 1$  such that for every  $x \in (0, c)$  and every graph  $G$  with  $\text{ind}_H(G) \leq (x^d|G|)^{|H|}$ , there is an  $x$ -sparse or  $x$ -dense  $(\ell(x), \lfloor x^d|G| \rfloor)$ -blockade in  $G$ . Also, for each integer  $s \geq 0$ ,  $\ell_s: (0, \frac{1}{2}) \rightarrow \mathbb{R}^+$  is the function defined by

$$\ell_s(x) := 2^{(\log \frac{1}{x})^{\frac{s}{s+1}}}$$

for all  $x \in (0, \frac{1}{2})$ . In this section we use the results of the previous two sections, together with 2.3, to prove 2.4, which we restate:

**5.1** *Every path  $P$  is  $\ell_s$ -divisive for all integers  $s \geq 0$ .*

**Proof.** The proof is by induction on  $s$ . For  $s = 0$ , the result is due to Fox and Sudakov [5], extending a theorem of Erdős and Hajnal [4]; indeed, they proved that every graph is  $\ell_0$ -divisive. So, we assume that  $s \geq 1$ , and  $P$  is  $\ell_{s-1}$ -divisive. By 2.3, with  $\ell = \ell_{s-1}$ , we deduce that there exists  $C > 0$  such that for every  $\varepsilon \in (0, \frac{1}{2})$ , if we define  $\delta > 0$  by

$$\log \frac{1}{\delta} = \frac{C(\log \frac{1}{\varepsilon})^2}{\log(\ell_{s-1}(\varepsilon))},$$

then for every graph  $G$  with  $\text{ind}_P(G) \leq (\delta|G|)^{|P|}$ , there is an  $\varepsilon$ -restricted  $S \subseteq V(G)$  with  $|S| \geq \delta|G|$ . But  $\log(\ell_{s-1}(\varepsilon)) = (\log \frac{1}{\varepsilon})^{\frac{s-1}{s}}$  and so

$$\log \frac{1}{\delta} = C \left( \log \frac{1}{\varepsilon} \right)^{\frac{s+1}{s}}.$$

Let  $b = 3^{\frac{s+1}{s}}C$ . We deduce:

(1) Let  $0 < x \leq 1/2$  and let  $y := 1/\ell_s(x)$ . Then for every graph  $G$  with  $\text{ind}_P(G) \leq (x^b|G|)^{|P|}$ , there is a  $y^3$ -restricted subset  $S \subseteq V(G)$  with  $|S| \geq x^b|G|$ .

Since  $x \leq 1/2$ , it follows that  $\ell_s(x) \geq 2$  and so  $y^3 \leq y \leq 1/2$ . By setting  $\varepsilon = y^3$  and

$$\log \frac{1}{\delta} = C \left( \log \frac{1}{y^3} \right)^{\frac{s+1}{s}} = b \log \frac{1}{x}$$

(that is,  $\delta = x^b$ ) we deduce that for every graph  $G$  with  $\text{ind}_P(G) \leq (x^b|G|)^{|P|}$ , there is a  $y^3$ -restricted  $S \subseteq V(G)$  with  $|S| \geq x^b|G|$ . This proves (1).

Now, let  $d = 3b|P| + b + 5$ , and choose  $c > 0$  with  $c \leq 1/2$ , and sufficiently small that

$$c^b \leq \frac{1}{\ell_s(c)} \leq \min \left( \frac{1}{2|P|}, \frac{1}{100} \right).$$

Let  $x \in (0, c)$  and let  $G$  be a graph with  $\text{ind}_P(G) \leq (x^d|G|)^{|P|}$ . We will show that there is an  $x$ -sparse or  $x$ -dense  $(\ell_s(x), \lfloor x^d|G| \rfloor)$ -blockade in  $G$ , and therefore that  $P$  is  $\ell_s$ -divisible. Suppose (for a contradiction) that there is no such blockade. As before, let  $y := 1/\ell_s(x)$ .

(2) For every  $S \subseteq V(G)$  with  $|S| \geq x^{d-b-5}|G|$ , there exists a  $(1 - y^3)$ -dense subset  $S' \subseteq S$  with  $|S'| \geq x^b|S|$ .

Suppose not. By (1) applied to  $G[S]$ , either  $\text{ind}_P(G[S]) > (x^b|S|)^{|P|}$ , or there is an  $y^3$ -sparse subset  $S' \subseteq S$  with  $|S'| \geq x^b|S|$ . In the first case,

$$\text{ind}_P(G) > (x^b|S|)^{|P|} \geq (x^d|G|)^{|P|}$$

(since  $x^b|S| \geq x^d|G|$ ), a contradiction. In the second case,  $|S'| \geq x^b|S|$ , and by 3.1 applied to  $G[S']$ , either

- $\text{ind}_P(G[S']) \geq (x^4|S'|)^{|P|}$ , or
- there is an  $x$ -sparse  $(1/y, \lfloor x^5|S'| \rfloor)$ -blockade in  $G[S']$ .

The first is impossible since  $x^4|S'| \geq x^{b+4}|S| \geq x^d|G|$ . If the second holds, then  $G$  admits an  $x$ -sparse  $(1/y, \lfloor x^{b+5}|G| \rfloor)$ -blockade and hence an  $x$ -sparse  $(1/y, \lfloor x^d|G| \rfloor)$ -blockade since  $d \geq b+5$ , again a contradiction. This proves (2).

In particular, (1) implies that for every  $S \subseteq V(G)$  with  $|S| \geq x^{3b|P|}|G|$ , there is a  $(1 - y^3)$ -dense subset  $S' \subseteq S$  with  $|S'| \geq x^b|S|$ , since  $x^{3b|P|} = x^{d-b-5}$ . We chose  $c$  such that  $\frac{1}{\ell_s(c)} \leq \frac{1}{100}$ , and since  $x < c$  and  $y = 1/\ell_s(x)$ , it follows that  $y \leq 1/100$ . By 4.1, with  $x$  replaced by  $x^b$ , we deduce that either:

- $\text{ind}_P(G) \geq (x^{3b|P|}|G|)^{|P|}$ ; or
- there is a  $(1 - x^b)$ -dense  $(1/y, \lfloor x^{3b|P|}|G| \rfloor)$ -blockade in  $G$ .

The first is impossible since  $(x^{3b|P|}|G|)^{|P|} \geq (x^d|G|)^{|P|}$  (because  $d \geq 3b|P|$ ). Thus there is a  $(1 - x^b)$ -dense, and hence  $(1 - x)$ -dense,  $(1/y, \lfloor x^d|G| \rfloor)$ -blockade in  $G$ . This proves 5.1.  $\blacksquare$

## 6 Concluding remarks

We have proved that all paths are near-viral; can we do the same for any other graphs? A *caterpillar* is a tree in which there is a path that contains at least one end of every edge of the tree; and the proof of this paper for paths can be adapted without difficulty to show that all caterpillars are near-viral. This leads naturally to asking whether all trees are near-viral, but we have not been able to prove that; we cannot handle the tree obtained from the 3-claw  $K_{1,3}$  by subdividing each edge once (this is the unique minimal tree that is not a caterpillar). But there are four further graphs (two complementary pairs) that we can prove are near-viral, all with six vertices and shown in the figure 3. The method of proof is similar. We omit all these proofs, which will appear in Tung Nguyen’s thesis [6].

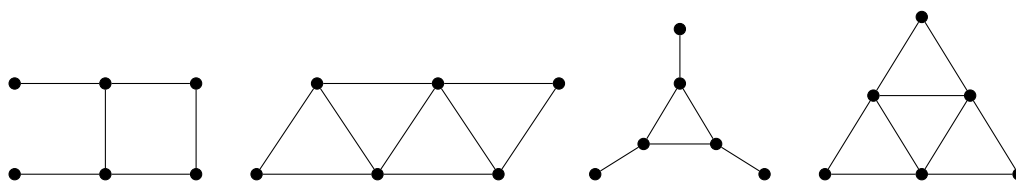


Figure 3: Four near-viral six-vertex graphs.

Another way to strengthen the current result is by looking at ordered graphs. An *ordered graph*  $G$  is a pair  $(G^{\natural}, \leq_G)$ , where  $G^{\natural}$  is a graph and  $\leq_G$  is a linear order of its vertex set. Induced subgraph containment for ordered graphs is defined in the natural way, respecting the orders of both graphs. A *zigzag path* is an ordered graph  $(G^{\natural}, \leq_G)$  where  $G^{\natural}$  is a path and the ordering is as in figure 4. The proof of this paper works for ordered graphs, with minor adjustments, showing that every zigzag path is near-viral (defining “near-viral” for ordered graphs in the natural way). We omit the details.

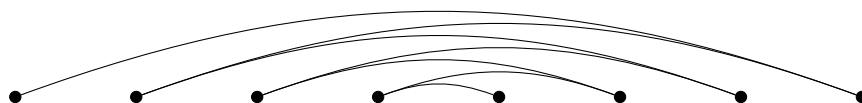


Figure 4: A zigzag path

## References

- [1] P. Blanco and M. Bucić, “Towards the Erdős-Hajnal conjecture for  $P_5$ -free graphs”, *Res. Math. Sci.*, **11**(2) (2024), [arXiv:2210.10755](https://arxiv.org/abs/2210.10755).
- [2] M. Bucić, T. Nguyen, A. Scott, and P. Seymour, “Induced subgraph density. I. A loglog step towards Erdős-Hajnal”, submitted for publication, [arXiv:2301.10147](https://arxiv.org/abs/2301.10147).
- [3] P. Erdős and A. Hajnal, “On spanned subgraphs of graphs”, *Contributions to Graph Theory and its Applications* (Internat. Colloq., Oberhof, 1977), 80–96, Tech. Hochschule Ilmenau, Ilmenau, 1977, [https://old.renyi.hu/~p\\_erdos/1977-19.pdf](https://old.renyi.hu/~p_erdos/1977-19.pdf)

- [4] P. Erdős and A. Hajnal, “Ramsey-type theorems”, *Discrete Applied Mathematics* **25** (1989), 37–52.
- [5] J. Fox and B. Sudakov, “Induced Ramsey-type theorems”, *Advances in Mathematics* **219** (2008), 1771–1800.
- [6] T. Nguyen, *Induced Subgraph Density*, PhD thesis, Princeton University, May 2025, in preparation.
- [7] T. Nguyen, A. Scott and P. Seymour, “Induced subgraph density. IV. New graphs with the Erdős-Hajnal property”, submitted for publication, [arXiv:2307.06455](https://arxiv.org/abs/2307.06455).
- [8] T. Nguyen, A. Scott and P. Seymour, “Induced subgraph density. VII. The five-vertex path”, in preparation, [arXiv:2312.15333](https://arxiv.org/abs/2312.15333).
- [9] V. Nikiforov, “Edge distribution of graphs with few copies of a given graph”, *Combin. Probab. Comput.* **15** (2006), 895–902.
- [10] V. Rödl, “On universality of graphs with uniformly distributed edges”, *Discrete Mathematics* **59** (1986), 125–134.