

# Sparse graphs without long induced paths\*

Oscar Defrain<sup>†</sup>      Jean-Florent Raymond<sup>‡</sup>

April 19, 2023

## Abstract

Graphs of bounded degeneracy are known to contain induced paths of order  $\Omega(\log \log n)$  when they contain a path of order  $n$ , as proved by Nešetřil and Ossona de Mendez (2012). In 2016 Esperet, Lemoine, and Maffray conjectured that this bound could be improved to  $\Omega((\log n)^c)$  for some constant  $c > 0$  depending on the degeneracy.

We disprove this conjecture by constructing, for arbitrarily large values of  $n$ , a graph that is 2-degenerate, has a path of order  $n$ , and where all induced paths have order  $O((\log \log n)^2)$ . We also show that the graphs we construct have linearly bounded coloring numbers.

**Keywords:** typical induced subgraphs, degeneracy, coloring numbers.

## 1 Introduction

In this paper logarithms are binary. For an integer  $k$ , a graph  $G$  is said to be  $k$ -*degenerate* if every subgraph of  $G$  has a vertex of degree at most  $k$ . The minimum  $k$  such that  $G$  is  $k$ -degenerate is called the *degeneracy* of  $G$ . Bounded degeneracy forms a very general notion of graph sparsity that includes classes of bounded treewidth, planar graphs, and more generally graph classes of bounded expansion. Because its definition sets constraints on all subgraphs, it is also used in numerous inductive proofs and greedy algorithms, for instance for graph coloring.

---

\*This work was supported by the ANR projects GRALMECO (ANR-21-CE48-0004) and DISTANCIA (ANR-17-CE40-0015).

<sup>†</sup>Aix Marseille Université, CNRS, LIS, Marseille, France

<sup>‡</sup>CNRS, LIMOS, Université Clermont-Auvergne, Clermont-Ferrand, France.

In the *Sparsity* textbook [NOdM12], Nešetřil and Ossona de Mendez proved that in graph classes of bounded degeneracy, paths and induced paths are tied in the following sense.

**Theorem 1.1** ([NOdM12]). *Let  $k \in \mathbb{N}$  and let  $G$  be a  $k$ -degenerate graph. If  $G$  has a path of order  $n$ , then  $G$  has an induced path of order at least*

$$\frac{\log \log n}{\log(k+1)}.$$

As every induced path is a path, the above result establishes a duality between long paths and induced paths for graphs of bounded degeneracy. Relating these objects is a very natural goal as they are among the most basic structures in graph theory. So it should not be a surprise that it was already investigated decades earlier, when Galvin, Rival, and Sands proved the following general result.

**Theorem 1.2** ([GRS82]). *There is a function  $f: \mathbb{N}^2 \rightarrow \mathbb{R}$  such that for every  $s$ , if  $G$  is a  $K_{s,s}$ -subgraph free graph that has a path of order  $n$ , then  $G$  has an induced path of order at least  $f(n, s)$ .*

This result can be related to the study of *unavoidable* (or *typical*) substructures of a graph class  $\mathcal{G}$  (i.e., substructures that can be found in every member of  $\mathcal{G}$ ), that originates from Ramsey's theorem and is currently receiving increasing attention when these structures are induced subgraphs.

If we consider graph classes that are closed under taking induced subgraphs, the above result describes the most general setting where such a function  $f$  could exist. Indeed, cliques and bicliques have paths that visit all their vertices, yet they do not have induced paths on more than 2 or 3 vertices, respectively. So any hereditary class where a result such as Theorem 1.2 holds should exclude arbitrarily large cliques and bicliques, and this amounts to excluding arbitrarily large biclique subgraphs. On the other hand, the function  $f$  defined in the proof of Theorem 1.2 is not optimal for many  $K_{s,s}$ -subgraph free graph classes (see below) and possibly not optimal neither for the statement of the theorem.

Relating the maximum orders of paths and induced paths is also of interest in the context of graph sparsity theory because of the following connection to the parameter treedepth. For a graph  $G$ , if we denote by  $\ell$  the maximum order of an induced path in it and by  $t$  its treedepth then

$$\log(\ell + 1) \leq t \leq \ell.$$

Therefore in graph classes where bounds such as that of Theorem 1.2 hold, treedepth is not only tied to long induced paths but also to long paths. We refer to Chapter 6

in [NOdM12] for a definition of treedepth, a proof of the above inequality, and a discussion on the links between induced paths and treedepth.

Following the proof of Theorem 1.1, the authors of *Sparsity* ask about the best possible bound.

**Question 1.3** ([NOdM12, Problem 6.1]). *For every  $k \in \mathbb{N}$ , what is the maximum function  $f_k: \mathbb{N} \rightarrow \mathbb{N}$  such that every  $k$ -degenerate graph  $G$  that has a path of order  $n$  has an induced path of order at least  $f_k(n)$ ?*

To the best of our knowledge, the current bounds on  $f_k$ , for  $k \geq 3$  and  $n \geq 1$ ,<sup>1</sup> are the following:

$$\frac{\log \log n}{\log(k+1)} \leq f_k(n) \leq k(\log n)^{2/k}.$$

The left-hand inequality is Theorem 1.1 and the right-hand bound, due to Esperet, Lemoine, and Maffray, is proved by the construction, for every  $t \geq 4$ , of an infinite sequence of chordal graphs of clique number  $t$  (hence  $(t-1)$ -degenerate) where the induced paths can be shown to be all short [ELM17].<sup>2</sup> The gap between the upper- and lower-bounds above is exponentially wide so it is indeed an interesting goal to investigate the order of magnitude of  $f_k$ .

Actually, for several classes of degenerate graphs the lower-bound can be drastically improved. A representative example is the following result of Esperet et al. stating that planar graphs contain induced paths of polylogarithmic order.

**Theorem 1.4** ([ELM17]). *There is a constant  $c$  such that if  $G$  is a planar graph and has a path of order  $n$ , then  $G$  has an induced path of order at least  $c\sqrt{\log n}$ .*

The result holds more generally for graph classes of bounded Euler genus, as shown in the same paper. This led its authors to conjecture, as an answer to Question 1.3, that polylogarithmic bounds could hold for graphs of bounded degeneracy.

**Conjecture 1.5** ([ELM17, Conjecture 1.1]). *For every  $k \in \mathbb{N}$  there is a constant  $c > 0$  such that if  $G$  is a  $k$ -degenerate graph and has a path of order  $n \geq 1$ , then  $G$  has an induced path of order at least  $(\log n)^c$ .*

Recently, Hilaire and the second author extended the catalog of graph classes known to satisfy Conjecture 1.5 with the following result.

---

<sup>1</sup>The left-hand bound also holds for  $k = 2$ . When  $k < 2$  we trivially have  $f_k(n) = n$ .

<sup>2</sup>We note that the analysis of the construction of [ELM17] cannot be improved much as such graphs are known to have induced paths of order  $\Omega((\log n)^{1/t})$ , as later proved in [HR23].

**Theorem 1.6** ([HR23]). *For every graph class  $\mathcal{G}$  that is closed under topological minors there is a constant  $c > 0$  such that if a graph  $G \in \mathcal{G}$  has a path of order  $n \geq 1$ , it has an induced path of order at least  $(\log n)^c$ .*

In this paper we disprove Conjecture 1.5 by proving the following statement.

**Theorem 1.7.** *There is a constant  $c$  such that for infinitely many integers  $n$ , there is a 2-degenerate graph  $G$  that has a path of order  $n$  and no induced path of order more than  $c(\log \log n)^2$ .*

Note that for every  $k \geq 2$ , every 2-degenerate graph is also  $k$ -degenerate. Hence Theorem 1.7 shows that  $f_k(n) = O((\log \log n)^2)$  and indeed disproves Conjecture 1.5 for every value of  $k$ .

A natural weakening of Conjecture 1.5 could be to consider graph classes of bounded coloring numbers. For every  $r \in \mathbb{N}_{\geq 1}$ , the  $r$ -coloring number is a graph invariant that can be seen as a distance- $r$  version of the degeneracy.<sup>3</sup> We say that a graph class has *bounded coloring numbers* if there is a function  $d: \mathbb{N} \rightarrow \mathbb{N}$  such that for every  $r \in \mathbb{N}_{\geq 1}$ , the  $r$ -coloring number of any  $G \in \mathcal{G}$  is at most  $d(r)$ . This notion plays an important role in the theory of graph sparsity due to its strong ties to the concept of bounded expansion.<sup>4</sup> In particular, as observed by Norin, if a graph class has coloring numbers bounded by a linear function, then this class has expansion bounded by a linear function (see [ER18, Observation 10]).

As the 1-coloring number is exactly the degeneracy, bounding the coloring numbers of a graph class is substantially stronger than merely bounding the degeneracy. The following shows that even with the requirement that the considered graphs come from a class of linearly bounded coloring numbers, Conjecture 1.5 does not hold. Let  $\mathcal{G}$  denote the class of the graphs from Theorem 1.7.

**Theorem 1.8.**  *$\mathcal{G}$  has coloring numbers bounded from above by a linear function.*

**Organization of the paper.** In the next section we introduce the basic terminology used in this paper. The construction is described in Section 3. Section 4 is devoted to the proofs of general properties on the graphs we construct, while in Section 5 we discuss the behavior of their induced paths. We finally prove Theorem 1.7 in Section 6. In Section 7 we prove Theorem 1.8. Directions for future research are given in Section 8.

---

<sup>3</sup>In this paper we only deal with the coloring numbers sometimes known as strong coloring numbers. See Section 7 for a formal definition.

<sup>4</sup>We will not use the concept of bounded expansion later in this paper so we refrain to define it and refer to [NOdM12] for a formal definition and more about its links with the coloring numbers.

## 2 Preliminaries

Unless stated otherwise we use standard graph theory terminology.

**Trees.** Our construction is defined starting from trees. To avoid ambiguity we use the synonym *node* when referring to the vertices of a tree. A *rooted tree* is a tree with a distinguished node called its *root*. The *leaves* are the nodes with degree one and different from the root. The other vertices, that are neither the root nor leaves, are called *internal nodes*. For every pair  $s, t$  of nodes of a rooted tree we write  $s \preceq t$  if  $s$  lies on the unique path connecting  $t$  to the root, and  $s \prec t$  if in addition  $s \neq t$ . We say that  $t$  is a *descendant* of  $s$  and that  $s$  is an *ancestor* of  $t$  whenever  $s \preceq t$ ;  $t$  is a *child* of  $s$  and  $s$  is the *parent* of  $t$  if in addition  $t$  and  $s$  are neighbors.

The *depth* of a node  $s$  in a rooted tree is the order of the unique path from  $s$  to the root, denoted  $\text{depth}(s)$ ; note in particular that the root has depth 1.

For every  $p \in \mathbb{N}$ , the *complete binary tree*  $T$  of depth  $p$  is the rooted tree defined as follows:

- if  $p = 1$  then  $T = K_1$ , rooted at its unique vertex;
- otherwise  $p > 1$  and  $T$  can be obtained from two disjoint copies of the complete binary tree of depth  $p - 1$  by adding a new vertex  $v$  adjacent to their roots and rooting  $T$  at  $v$ .

**Remark 2.1.** *The complete binary tree of depth  $p$  has order  $2^p - 1$ .*

## 3 The construction

We shall describe, for arbitrarily large  $n$ , the construction of a 2-degenerate graph with a path of order  $n$ —which will in fact be Hamiltonian—but no induced path of order  $c(\log \log n)^2$  for some constant  $c$ .

**Tree blow-up.** The *blow-up*  $H$  of a complete binary tree  $T$  is the graph obtained from  $T$  as follows (see also Figure 1):

- for every node  $s$  of  $T$  we create a clique  $K^s$  on 3 vertices and choose an injection  $\mu_s: N_T(s) \rightarrow E(K^s)$ ;
- for every edge  $st$  in  $T$  with  $s \preceq t$  we add to  $H$  a first path  $L(s, t) = u_1x_1y_1v_1$  on four vertices connecting an endpoint  $u_1$  of  $\mu_s(t)$  to an endpoint  $v_1$  of  $\mu_t(s)$ , and

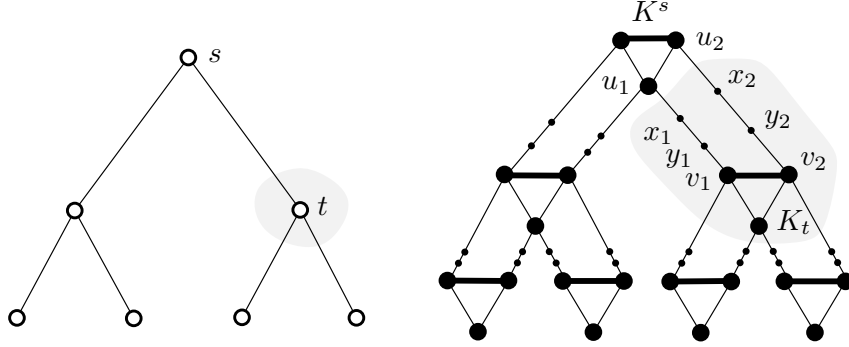


Figure 1: A complete binary tree of height 3 (left) and its blow-up (right) where large vertices represent cliques vertices, small ones represent subdivision vertices, and the bold edges represent top edges. In the gray area, the bag of  $t$ , i.e., the set of vertices that originated from  $t$  in the construction of the blow-up.

a second path  $R(s, t) = u_2x_2y_2v_2$  on four vertices connecting the other endpoint  $u_2$  of  $\mu_s(t)$  to the other endpoint  $v_2$  of  $\mu_t(s)$ ; we stress the fact that the endpoints of  $\mu_s(t)$  and  $\mu_t(s)$  are chosen arbitrarily, and that vertices  $x_1, x_2, y_1, y_2$  are added during the construction.

In the definition above we call *top-predecessors* of the node  $t$  the vertices in the set  $\mathbf{tpred}(t) = \{x_1, x_2\}$ , *bottom-predecessors* of  $t$  the vertices in  $\mathbf{bpred}(t) = \{y_1, y_2\}$ , and shall note  $\mathbf{pred}(t)$  the union of these two sets. We call *bag* of node  $t$  the set  $V(K^t) \cup \mathbf{pred}(t)$ . The edges (respectively vertices) of  $H$  that belong to  $K^s$  for some node  $s \in V(T)$  are called *clique edges* (respectively *clique vertices*), while the others are called *tree edges* of  $H$  (respectively *subdivision vertices*).<sup>5</sup> For a node  $t \in V(T)$  with parent  $s$ , the *top edge* of the clique  $K^t$  in  $H$  is the edge  $\mu_t(s)$ . The *top edge* of the root  $r$  of  $T$  is the only edge of  $K^r$  that has no preimage by  $\mu_r$ , and it shall also be referred to as the *root edge* of  $H$ .

**Nested intervals systems.** Let  $h: \mathbb{N}_{\geq 1} \rightarrow \mathbb{N}$  be the function defined for every  $\ell \in \mathbb{N}_{\geq 1}$  by the following recurrence relation:

$$\begin{cases} h(1) = 3 \\ h(\ell) = 2 + 2 \cdot h(\ell - 1) \end{cases} \quad \text{for every } \ell > 1. \quad (1)$$

<sup>5</sup>Subdivision vertices are used later in our construction in order to achieve low degeneracy. It seems possible to avoid them if we only aim for a 4-degenerate graph.

Observe that  $h(\ell)$  may as well be computed by the following explicit formula for every  $\ell \in \mathbb{N}_{\geq 1}$ :

$$h(\ell) = 5 \cdot 2^{\ell-1} - 2. \quad (2)$$

If  $X$  is a set of pairs of integers and  $i \in \mathbb{N}$ , we denote by  $X \oplus i$  the set  $\{(x+i, x'+i) : (x, x') \in X\}$ . For every  $\ell \in \mathbb{N}_{\geq 1}$  the set  $\mathcal{N}_\ell$  is recursively defined as follows:

$$\begin{cases} \mathcal{N}_1 = \{(1, 3)\}, \text{ and} \\ \mathcal{N}_\ell = \{(1, h(\ell))\} \cup (\mathcal{N}_{\ell-1} \oplus 1) \cup (\mathcal{N}_{\ell-1} \oplus (h(\ell-1) + 1)) \text{ if } \ell > 1. \end{cases}$$

The elements of  $\mathcal{N}_\ell$  are called *intervals*. Intuitively  $\mathcal{N}_\ell$  is obtained by taking two copies of  $\mathcal{N}_{\ell-1}$  (appropriately shifted so that they start after integer 1 and do not intersect) and adding a new interval  $(1, h(\ell))$  containing the two copies. See Figure 2 for an illustration.

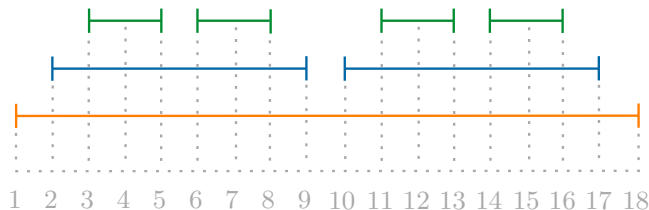


Figure 2: The intervals of  $\mathcal{N}_3$  with intervals of rank 1, 2, and 3 depicted from top to bottom in green, blue, and orange, respectively.

The following easy properties of  $\mathcal{N}_\ell$  can be proved by a straightforward induction:

**Remark 3.1.** For every  $\ell \in \mathbb{N}_{\geq 1}$  the following holds:

1. the endpoints of the intervals in  $\mathcal{N}_\ell$  range from 1 to  $h(\ell)$ ;
2. the intervals of  $\mathcal{N}_\ell$  all have distinct endpoints;
3. every interval of  $\mathcal{N}_\ell$  is of the form  $(i, i + h(a) - 1)$  for some  $i \in \{1, \dots, h(\ell)\}$  and  $a \in \{1, \dots, \ell\}$ ;
4. for every interval  $(i, j)$  in  $\mathcal{N}_\ell$  there is no other interval  $(i', j')$  in  $\mathcal{N}_\ell$  such that  $i < i' < j < j'$  (informally, intervals do not cross).

We call *rank* of an interval  $(i, j) \in \mathcal{N}_\ell$  the aforementioned integer  $a \in \{1, \dots, \ell\}$  such that  $j = i + h(a) - 1$ .

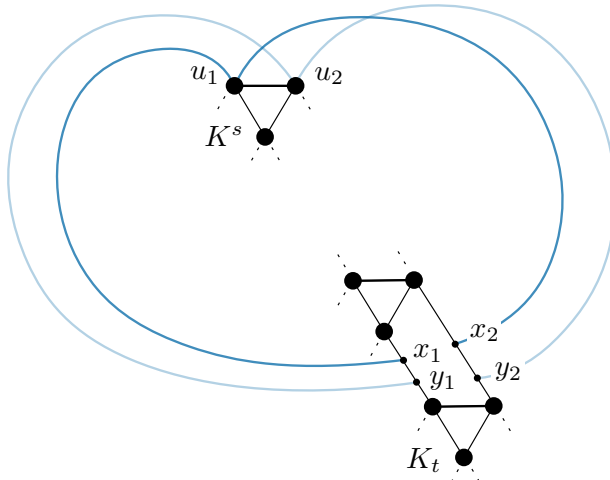


Figure 3: The edges in the set  $\text{ribs}_\ell(s, t)$  for two nodes  $s, t \in V(B_\ell)$  such that  $s \prec t$ .

**Ribs.** For every  $\ell \in \mathbb{N}_{\geq 1}$  we denote by  $B_\ell$  the complete binary tree of depth  $h(\ell)$  and by  $H_\ell$  its blow-up. The next construction is illustrated in Figure 3. Let  $\text{ribs}_\ell$  be the function defined on every pair  $(s, t) \in V(B_\ell)^2$  of nodes such that  $s \prec t$  as the following set of edges over  $V(H_\ell)$ :<sup>6</sup>

$$\text{ribs}_\ell(s, t) = \{u_1x_1, u_1x_2, u_2y_1, u_2y_2\},$$

where  $\{x_1, x_2\} = \text{tpred}(t)$ ,  $\{y_1, y_2\} = \text{bpred}(t)$ , and  $u_1u_2$  is the top edge of  $K^s$ . Note that such a function is not unique,<sup>7</sup> however any will suffice for our purpose; see Figure 3 for an example. The graph  $G_\ell$  is obtained from  $H_\ell$  after the addition of the set of edges  $\text{ribs}_\ell(s, t)$  for every pair of nodes  $s, t \in V(B_\ell)$  of respective depth  $i, j$  such that  $s \prec t$  and  $(i, j) \in \mathcal{N}_\ell$ . We call an edge  $uv$  in that set a *rib*. Hence the edges of  $G_\ell$  are partitioned into tree edges, clique edges, and ribs; see Figure 4 for a representation of  $G_\ell$ .

<sup>6</sup>Note that these edges do not exist in  $H_\ell$ . We define this set in order to later construct a graph by adding these edges to  $H_\ell$ .

<sup>7</sup>It depends on which vertex in  $\text{tpred}(t)$  we decided to call  $x_1$  in the construction of the blow-up, and similarly for  $\text{bpred}(t)$  and the top edge of  $s$ .



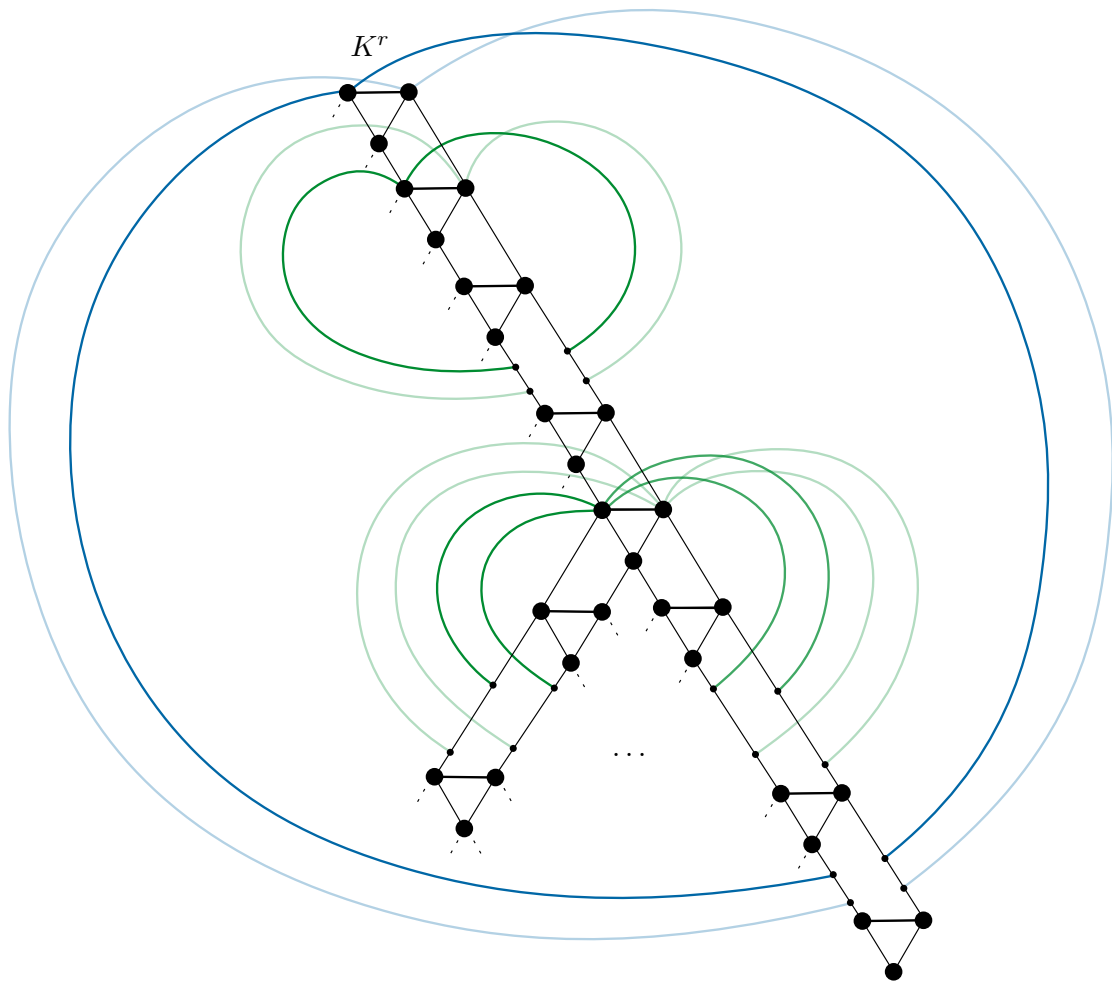


Figure 4: A partial representation of the graph  $G_2$  with root clique  $K^r$ . For readability, subdivision vertices that are not incident to ribs are omitted. Ribs corresponding to intervals of rank 1 and 2 are depicted in green and blue, respectively.

Because much of the structure of  $G_\ell$  is inherited from  $B_\ell$ , some properties of vertices of  $G_\ell$  will be conveniently defined via this tree. For this purpose we define  $\pi: V(G_\ell) \rightarrow V(B_\ell)$  as the function mapping every clique vertex  $u$  of  $G_\ell$  to the corresponding node  $s$  in  $B_\ell$  (i.e., the unique node  $s$  such that  $u \in V(K^s)$ ) and every subdivision vertex  $v$  of  $G_\ell$  to the unique node  $t$  in  $B_\ell$  such that  $v \in \text{pred}(t)$ . We define the depth  $\text{depth}(u)$  of a vertex  $u \in V(G_\ell)$  as the depth of its corresponding node  $\pi(u)$  in  $B_\ell$ . The following stems from the definition of a blow-up.

**Remark 3.2.** *The set  $\{\pi^{-1}(t) : t \in V(B_\ell)\}$  defines a partition of  $V(G_\ell)$  with  $|\pi^{-1}(t)| = 3$  when  $t$  is the root of  $B_\ell$ , and  $|\pi^{-1}(t)| = 7$  otherwise.*

Note that  $\pi^{-1}(t)$  is exactly denotes the bag of  $t$ ; see Figure 1 for an example.

## 4 The properties of $G_\ell$

A *Hamiltonian path* in a graph is a path that visits all its vertices.

**Lemma 4.1.** *For every integer  $\ell \geq 1$ , the graph  $G_\ell$  has a Hamiltonian path.*

*Proof.* In this proof we denote by  $F_\ell$  the blow-up of the complete binary tree of depth  $\ell$ , for every  $\ell \in \mathbb{N}_{\geq 1}$ . We will show the following.

*For every integer  $\ell \geq 1$ ,  $F_\ell$  has a Hamiltonian path linking the two endpoints of its root edge.*

Note that the above immediately implies the desired statement as  $H_\ell = F_{h(\ell)}$  is a spanning subgraph of  $G_\ell$ . The proof is by induction on  $\ell$ . In the case  $\ell = 1$ ,  $F_\ell$  is the triangle and the statement clearly holds.

Let us assume from now on that  $\ell > 1$  and that the property holds for  $\ell - 1$ . It follows from the definition that  $F_\ell$  can be constructed from a triangle  $z_1 z_2 z_3$  (where  $z_1 z_2$  will be the root) and two disjoint copies  $F, F'$  of  $F_{\ell-1}$  of respective root edges  $xy$  and  $x'y'$  by adding paths  $P(u, v)$  joining  $u$  to  $v$  for every  $(u, v) \in \{(z_1, x), (y, z_3), (z_3, x'), (y', z_2)\}$ . Then a Hamiltonian path  $P$  of  $F_\ell$  from  $z_1$  to  $z_2$  can be constructed as follows: from  $z_1$  we follow  $P(z_1, x)$ , then the Hamiltonian path of  $F$  between  $x$  and  $y$  (given by the induction hypothesis) leads us to  $y$  and we reach  $z_3$  via  $P(y, z_3)$ . The second half of the path through  $F'$  up to  $z_2$  is symmetric. This proves the above statement, and hence the lemma.  $\square$

**Lemma 4.2.** *For every integer  $\ell \geq 1$ ,  $|V(G_\ell)| \geq 2^{2^{\ell+1}}$ .*

*Proof.* Let  $r$  denote the root of  $B_\ell$ . As observed in Remark 3.2, the vertex set of  $G_\ell$  can be partitioned into sets  $\pi^{-1}(s)$  of size either 3 if  $s$  is the root, or 7 otherwise. Hence

$$\begin{aligned} |V(G_\ell)| &= 7 \cdot |B_\ell| - 4 \\ &= 7 \cdot 2^{5 \cdot 2^{\ell-1} - 2} - 11 && \text{(using Remark 2.1 and (2))} \\ &\geq 2^{2^{\ell+1}}. \end{aligned}$$

$\square$

**Lemma 4.3.** *For every integer  $\ell \geq 1$ , the graph  $G_\ell$  is 2-degenerate.*

*Proof.* Let  $G$  be a (non-empty) induced subgraph of  $G_\ell$  and let  $D$  denote the set of vertices of maximal depth  $d$  in  $G$ . We define 4 subsets of  $D$  as follows:

- $D_1$  is the set of clique vertices in  $D$  that are not incident with the top edge of their clique;
- $D_2$  is the set of clique vertices in  $D$  that are incident with the top edge of their clique;
- $D_3$  is the set of bottom-predecessors in  $D$ ;
- $D_4$  is the set of top-predecessors in  $D$ .

Clearly this forms a partition of  $D$  and one of these sets is not empty. Let  $i$  be the minimum index such that  $D_i$  is not empty and let  $v \in D_i$ . Let us show that  $\deg_G(v) = 2$ . We distinguish cases according to the value of  $i$ :

Case  $i = 1$ : In  $G_\ell$ , the vertex  $v$  has  $\leq 4$  neighbors: the two other elements of  $K^{\pi(v)}$  and at most two vertices  $x, y \in \text{tpred}(s) \cup \text{tpred}(t)$  for  $s$  and  $t$  the children of  $\pi(v)$  in  $B_\ell$ , if any. As  $\text{depth}(s) = \text{depth}(t) > d$  and by maximality of  $d$ , none of  $x$  and  $y$  belongs to  $G$ . Hence  $v$  has at most two neighbors in this graph.

Case  $i = 2$ : Recall that if  $uu'$  is a rib of  $G_\ell$  and  $u$  is a clique vertex, then  $\text{depth}(u) < \text{depth}(u')$ . Therefore, as  $v$  is a clique vertex and  $d$  is maximum, no rib of  $G$  is adjacent to  $v$ . Besides, at most one other vertex of  $K^{\pi(v)}$  belongs to  $G$ , as otherwise  $D_1$  would not be empty. Finally, only one other vertex is adjacent to  $v$  in  $G_\ell$  (an element of  $\text{bpred}(\pi(v))$ ), so  $\deg_G(v) \leq 2$ .

Case  $i = 3$ : A vertex  $u \in \text{bpred}(s)$  for some  $s \in V(B_\ell)$  has the following neighbors in  $G_\ell$ : a vertex of the top edge of  $K^s$ , a vertex of  $\text{tpred}(s)$  and possibly a third neighbor connected via a rib (see Remark 3.1). The first type of neighbor does not exist in  $G$  as  $D_2 = \emptyset$ , so here again  $\deg_G(v) \leq 2$ .

Case  $i = 4$ : Similarly as above, a vertex  $u \in \text{tpred}(s)$  for some  $s \in V(B_\ell)$  has the following neighbors in  $G_\ell$ : a vertex of  $\text{bpred}(s)$ , a vertex of  $K^t$  (where  $t$  denotes the parent of  $s$  in  $B_\ell$ ) and possibly a neighbor connected via a rib. The first type of neighbor does not exist in  $G$  as  $D_3 = \emptyset$ , and again  $\deg_G(v) \leq 2$ .

We proved that every subgraph of  $G$  contains a vertex of degree at most 2. This shows that  $G_\ell$  is indeed 2-degenerate.  $\square$

## 5 Ribs, sources, and their properties

In the rest of the paper we fix  $\ell \in \mathbb{N}_{\geq 1}$ . A node  $s$  of  $B_\ell$  is a *source* if there is an interval  $(i, j) \in \mathcal{N}_\ell$  such that  $s$  has depth  $i$ . Intuitively this means that in  $G_\ell$  the clique  $K^s$  will send ribs to vertices deeper (i.e., in  $\text{pred}(t)$  for the descendants  $t$  of  $s$  of depth  $j$ ) in  $G_\ell$ . For a source  $s$  of  $B_\ell$  the *rank* of  $s$  is defined as the rank of  $(i, j)$ , i.e., the integer  $a \in \{1, \dots, \ell\}$  such that  $j - i + 1 = h(a)$ . We denote by  $B_\ell(s)$  the subtree of  $B_\ell$  rooted at  $s$  and of depth  $h(a)$ . This means that the leaves of  $B_\ell(s)$  are exactly those vertices  $t$  such that in  $G_\ell$   $K_s$  sends ribs to the predecessors of  $t$ . The graph  $G_\ell(s)$  is defined as the subgraph of  $G_\ell$  induced by  $\pi^{-1}(V(B_\ell(s)))$ .

Let  $Q$  be an induced path of  $G_\ell$ . The following definition is crucial in the rest of the proof: a source  $s$  is said to be *Q-special* if  $Q$  has two vertices  $u, v$  such that  $u$  is part of the top edge of  $K^s$  and  $\pi(v)$  is an internal node of  $B_\ell(s)$ . For every vertex  $v$  of  $G_\ell$ , we define  $\tau(v)$  as the minimum rank of a source  $s$  such that  $\pi(v)$  is an internal node of  $B_\ell(s)$ . Notice that if  $s$  is the root or a leaf of  $B_\ell$  then  $\tau$  is not defined: we set  $\tau(v) = \ell + 1$  in this case.

**Remark 5.1.** *Let  $v \in V(G_\ell)$  and let  $s$  be a source of  $B_\ell$  of rank  $a \in \{1, \dots, \ell\}$ .*

1. *if  $\pi(v)$  is an internal node of  $B_\ell(s)$  then  $\tau(v) \leq a$ ;*
2. *if  $\pi(v) = s$  or  $\pi(v)$  is a leaf of  $B_\ell(s)$ , then  $\tau(v) = a + 1$ .*

In a graph  $G$ , we say that a set  $X \subseteq V(G)$  *separates* two sets  $Y, Z \subseteq V(G)$  if every path from a vertex of  $Y$  to a vertex of  $Z$  intersects  $X$ .

**Lemma 5.2.** *Let  $s$  be a source of  $B_\ell$ , let  $L$  be the set of leaves of  $B_\ell(s)$ . Then*

$$X_s = V(K^s) \cup \bigcup_{t \in L} \pi^{-1}(t)$$

*separates  $V(G_\ell(s)) \setminus \text{pred}(s)$  from the other vertices of  $G_\ell$ .*

*Proof.* The node  $s$  is a source so there is an interval  $(i, j) \in \mathcal{N}_\ell$  such that  $\text{depth}(s) = i$  and  $\text{depth}(t) = j$  for every leaf  $t$  of  $B_\ell(s)$ . By Remark 3.1 there is no other interval  $(i', j') \in \mathcal{N}_\ell$  such that  $i \leq i' \leq j \leq j'$  or  $i' \leq i \leq j' \leq j$  so in the construction of  $G_\ell$  there is no rib with only one endpoint in  $G_\ell(s)$ . Therefore every edge leaving  $V(G_\ell(s)) \setminus \text{pred}(s)$  (i.e., with only one endpoint in that set) is a tree edge. Let  $v \in \pi^{-1}(t)$  for some internal node of  $B_\ell(s)$ . By construction every vertex that is connected to  $v$  via a tree edge is either in the bag of  $t$  or in the bag of a neighbor of  $t$  (possibly  $s$  or a leaf of  $B_\ell(s)$ ). In particular,  $v$  has no neighbor in  $(V(G_\ell) \setminus V(G_\ell(s))) \cup \text{pred}(s)$ . This shows that a path from  $v$  to  $(V(G_\ell) \setminus V(G_\ell(s))) \cup \text{pred}(s)$  goes through the set  $X_s$ , as claimed.  $\square$

We remark the following as a consequence of the definition of depth.

**Remark 5.3.** *If  $uv$  is an edge of  $G_\ell$  such that  $\text{depth}(u) < \text{depth}(v)$  then either  $uv$  is a tree edge and  $v$  a top-predecessor, or it is a rib of source  $\pi(u)$ .*

**Lemma 5.4.** *If  $uv$  is an edge of  $G_\ell$  such that  $\tau(u) > \tau(v)$  then  $uv$  is a tree edge and  $\tau(u) = \tau(v) + 1$ . If in addition  $\text{depth}(v) \geq \text{depth}(u)$ , then  $\pi(u)$  is a source and  $u \in V(K^{\pi(u)})$ .*

*Proof.* Let  $s \in V(B_\ell)$  be a source of minimum rank such that  $\pi(v)$  is an internal node of  $B_\ell(s)$ . Such a node exists as  $\tau(v) < \tau(u) \leq \ell + 1$  so  $\tau(v) \leq \ell$ , hence  $\pi(v)$  is in particular an internal node of  $B_\ell(r)$  where  $r$  is the root of  $B_\ell$ . By definition the rank of  $s$  is  $\tau(v)$ .

Observe that  $\pi(u)$  cannot be an internal node of  $B_\ell(s)$  as otherwise by Remark 5.1 we would have  $\tau(u) \leq \tau(v)$ . As proved in Lemma 5.2,  $X_s$  (as defined in the statement of that lemma) separates the vertices  $w$  such that  $\pi(w)$  is an internal node of  $B_\ell(s)$  from the rest of the graph, hence  $u \in X_s$ . By Remark 5.1,  $\tau(u) = \tau(v) + 1$  and we derive that  $uv$  is a tree edge.

In the case where  $\text{depth}(u) \leq \text{depth}(v)$ , from  $u \in X_s$  we deduce  $\pi(u) = s$  (hence it is a source). Since only vertices of  $K^s$  have neighbors in bags of internal nodes of  $B_\ell(s)$ , we furthermore have  $u \in V(K^s)$ .  $\square$

**Lemma 5.5.** *Let  $Q$  be an induced path of  $G_\ell$ . Then there is a unique node  $t \in V(B_\ell)$  of minimum depth subject to  $\pi^{-1}(t) \cap V(Q) \neq \emptyset$ .*

*Proof.* Let us assume towards a contradiction that there are two such nodes  $t, t'$ . As they have the same depth, they are not comparable in  $B_\ell$ . Besides, observe that every edge of  $G_\ell$  connects vertices whose image by  $\pi$  is  $\preceq$ -comparable. Therefore the subpath of  $Q$  linking the bag of  $t$  to that of  $t'$  contains a vertex of the bag of  $t''$  for some common ancestor  $t''$  of  $t$  and  $t'$ , which contradicts the minimality of the depth of those vertices.  $\square$

**Lemma 5.6.** *There is a constant  $c_{5.6}$  such that if  $Q = u_1 \dots u_q$  is an induced path of  $G_\ell$  with  $\tau(u_i) = \tau(u_1)$  for all  $2 \leq i \leq q$ , then  $|Q| \leq c_{5.6}$ .*

*Proof.* Let  $Q$  be such an induced path and  $s$  be the source of rank  $a = \tau(u_1)$  such that  $\pi(u_1)$  is an internal node of  $B_\ell(s)$ . By Remark 5.1 we have  $\tau(v) = a + 1$  if  $\pi(v) = s$  or  $\pi(v)$  is a leaf of  $B_\ell(s)$ . By Lemma 5.2 this implies that  $Q$  is contained in the union of the bags of the internal nodes of  $B_\ell(s)$ . Let us call this union  $Z$ . If  $a = 1$  there are at most two such nodes so clearly  $Q$  has bounded size. So we may now assume  $a > 1$ .

As  $s$  is a source, there is an interval  $(i, i'') \in \mathcal{N}_\ell$  such that  $\text{depth}(s) = i$  and  $\text{depth}(t) = i''$  for every leaf  $t$  of  $B_\ell(s)$ . We call  $s_1$  and  $s_2$  the two children of  $s$ . Let  $i' = i + h(a - 1)$ . By construction  $(i + 1, i'), (i' + 1, i'' - 1) \in \mathcal{N}_\ell$ ; see Figure 5 for a representation of  $B_\ell(s)$  by depths  $i, i + 1, \dots, i''$ . (Notice that  $i'' \geq i + 2$ , by the third item of Remark 3.1, the definition of the function  $h$  and the fact that  $a > 1$ .) Let  $D$  be the set of descendants of  $s$  that have depth  $i' + 1$  in  $G_\ell$ .

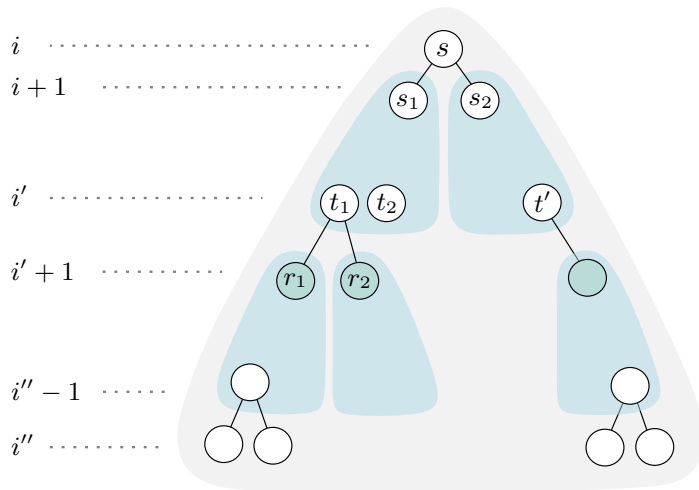


Figure 5: The situation of Lemma 5.6; full nodes represent elements in  $D$ .

Let  $r \in D$  and let  $t$  be a leaf of  $B_\ell(r)$ . Then  $t$  lies at depth  $i'' - 1$  in  $B_\ell$ . Observe that in  $G_\ell$  each edge incident to the bag of  $t$  is of one of the following types:

- a tree edge to a vertex  $v \in \pi^{-1}(t')$  where  $t'$  is the parent or a child of  $t$ ;
- a rib to the top edge of  $K^r$ .

As observed above, edges of the former type lead to a vertex  $v$  with  $\tau(v) \neq a$ . Therefore if  $Q$  visits the bag of  $t$  and other vertices of  $Z$  then  $Q$  follows a rib to the top edge of  $K^r$ . As  $Q$  may visit the top edge of  $K^r$  at most twice we deduce that there are at most three nodes of  $B_\ell(r)$  whose bag is intersected by  $Q$  (those of  $r$  and of two leaves of  $B_\ell(r)$ ).

For each  $j \in \{1, 2\}$  the same argument applies to  $s_j$  and the leaves of  $B_\ell(s_j)$ : there are at most three nodes of  $B_\ell(s_j)$  whose bags are intersected by  $Q$ , which are  $s_j$  and at most 2 leaves. We call these leaves  $t, t'$  (and choose them arbitrarily if  $Q$  visits less than two bags of leaves).

Each of  $t$  and  $t'$  has two children in  $B_\ell$ . For each such child  $r$ , we observed above (as  $r \in D$ ) that  $Q$  intersects at most 3 bags of nodes of  $B_\ell(r)$ . This shows that among the bags of descendants of  $s_j$ ,  $Q$  intersects at most 15 of them. So in total  $Q$  is contained in the union of at most 30 bags. The bound then follows from the fact that bags have bounded size (Remark 3.2).  $\square$

## 6 The induced paths of $G_\ell$ are short

An induced path  $v_1 \dots v_q$  of  $G_\ell$  is called *monotone* if  $\text{depth}(v_i) \leq \text{depth}(v_j)$  for all  $1 \leq i < j \leq q$ . It is called *bimonotone* if there exists  $k \in \{1, \dots, q\}$  such that  $v_k \dots v_1$  and  $v_k \dots v_q$  are monotone.

**Lemma 6.1.** *There is a constant  $c_{6.1}$  such that the following holds. Let  $Q$  be an induced path in  $G_\ell$  such that no source in  $B_\ell$  is  $Q$ -special. Then  $|Q| \leq c_{6.1} \cdot \ell$ .*

*Proof.* Let  $Q = v_1 \dots v_q$  be such an induced path. The proof is split into three claims.

**Claim 6.2.** *The path  $Q$  is bimonotone*

*Proof.* By Lemma 5.5 there is a unique node  $t \in V(B_\ell)$  minimizing the depth and whose bag is visited by  $Q$ . Let  $k$  be the minimum index such that  $v_k$  belongs to the bag of  $t$ . We now show that  $\text{depth}(v_k) \leq \dots \leq \text{depth}(v_1)$ , i.e., that  $v_k \dots v_1$  is monotone. Let us assume toward a contradiction that it is not the case, and let  $i \in \{1, \dots, k\}$  be the maximum integer such that  $\text{depth}(v_i) > \text{depth}(v_{i-1})$ . Take  $j \in \{i, \dots, k\}$  minimum such that  $\text{depth}(v_{j+1}) < \text{depth}(v_i)$ . By the choice of  $i, j$  we note that the vertices in the subpath  $v_i \dots v_j$  have equal depth. Because the edges in  $G_\ell$  only connect vertices of bags corresponding to  $\preceq$ -comparable vertices of  $B_\ell$ , two adjacent vertices of equal depth necessarily belong to the same bag. Therefore,  $\{v_i, \dots, v_j\}$  is a subset of the bag of a single node  $s \in V(B_\ell)$ . Furthermore by Remark 5.3 the edges  $v_{i-1}v_i$  and  $v_jv_{j+1}$  are either tree edges connecting a clique vertex to a top-predecessor, or ribs.

We distinguish cases depending on the type of the edges  $v_{i-1}v_i$  and  $v_jv_{j+1}$ .

In the case where both  $v_{i-1}v_i$  and  $v_jv_{j+1}$  are tree edges,  $v_{i-1}$  and  $v_{j+1}$  belong to the same clique  $K^{t'}$  (where  $t'$  is the parent of  $t$  in  $B_\ell$ ), which contradicts the fact that  $Q$  is induced. We now distinguish cases depending on whether  $v_{i-1}v_i$  or  $v_jv_{j+1}$  is a rib.

In the case where  $v_{i-1}v_i$  is a rib with  $v_{i-1}$  in the top edge of some source  $s \in V(B_\ell)$ , we consider the edge  $v_jv_{j+1}$ . Note that it cannot be a rib as otherwise  $v_{i-1}$  and  $v_{j+1}$  would both be part of the top edge of  $K^s$ , contradicting the fact that  $Q$  is induced.

So it is a tree edge. As  $\text{depth}(v_{j+1}) = \text{depth}(v_i) - 1$ ,  $\pi(v_{j+1})$  is an internal node of  $B_\ell(s)$ . Hence  $s$  is  $Q$ -special, a contradiction. The remaining case where  $v_j v_{j+1}$  is a rib is symmetric. This proves that such a  $i$  does not exist. The proof of the monotonicity of  $v_k \dots v_q$  is symmetric, yielding the bimonotonicity of  $Q$ .  $\lrcorner$

By the virtue of Claim 6.2 there is an integer  $k \in \{2, \dots, q\}$  such that each of  $v_k \dots v_1$  and  $v_k \dots v_q$  is monotone. We choose such a  $k$  minimum and set  $k'$  as the maximum integer such that  $v_k, \dots, v_{k'}$  all lie in a same bag and  $v_{k'} \dots v_q$  is monotone, with possibly  $k = k'$ .

**Claim 6.3.** *For every  $i, j \in \{k' + 1, \dots, q\}$  with  $i < j$ , we have  $\tau(v_i) \leq \tau(v_j)$ .*

*Proof.* The proof is by contradiction. Let  $i \in \{k' + 1, \dots, q - 1\}$  be the minimum integer such that  $\tau(v_{i+1}) < \tau(v_i)$  and let  $s = \pi(v_i)$ . By Lemma 5.4 and since  $\text{depth}(v_i) \leq \text{depth}(v_{i+1})$ ,  $s$  is a source,  $v_i$  belongs to the clique  $K^s$ ,  $v_i v_{i+1}$  is a tree edge and we get that  $v_{i+1}$  belongs to the bag of a child of  $s$ . If  $v_i$  is a vertex incident to the top edge of  $K^s$  then  $s$  is  $Q$ -special, a contradiction. Thus  $v_i$  is the only vertex of  $K^s$  that does not belong to its top edge. By monotonicity of  $v_{k'} \dots v_q$  we have  $\text{depth}(v_{k'}) \leq \text{depth}(v_{i-1}) \leq \text{depth}(v_i)$  and so  $v_{i-1}$  cannot be a top-predecessor of a child of  $s$ . Hence  $v_{i-1}$  must belong to the top edge of  $K^s$ , yielding that  $s$  is  $Q$ -special, a contradiction.  $\lrcorner$

A symmetric proof shows the following symmetric statement.

**Claim 6.4.** *For every  $i, j \in \{1, \dots, k - 1\}$  with  $i < j$ , we have  $\tau(v_i) \geq \tau(v_j)$ .*

We are now ready to conclude the proof. Recall that the function  $\tau$  has values in  $\{1, \dots, \ell + 1\}$ . It is non-increasing on  $v_1 \dots v_{k-1}$  (Claim 6.4) and has values in  $\{1, \dots, \ell\}$  (as  $\tau(v_{k-1}) < \tau(v_k)$ , by minimality of  $k$ ), non-decreasing on  $v_{k'+1} \dots v_q$  (Claim 6.3) and does not keep the same value on more than  $c_{5.6}$  consecutive vertices (Lemma 5.6). Now since  $v_k \dots v_{k'}$  lies in a single bag, it has order  $\leq 7$ . We conclude that  $Q$  has order at most  $c_{5.6}(2\ell + 1) + 7$ , which is bounded from above by  $c \cdot \ell$  for some constant  $c$ , as  $\ell \geq 1$ .  $\square$

We now show that if an induced path visits the top edge of the bag of a source  $s$ , as well as a bag of an internal node of  $B_\ell(s)$ , then one of its sides lives in the vertices of the bags of  $B_\ell(s)$ .

**Lemma 6.5.** *Let  $Q = v_1 \dots v_q$  be an induced path of  $G_\ell$  with  $v_1$  an endpoint of the top edge of  $K^s$  for some source  $s$  in  $B_\ell$ . If  $s$  is a  $Q$ -special source, then  $V(Q) \subseteq V(G_\ell(s)) \setminus \text{pred}(s)$ .*



*Proof.* Let  $X_s$  be the set defined in Lemma 5.2. Let  $i$  be the minimum integer such that  $\pi(v_i)$  is an internal node of  $B_\ell(s)$ . Such an integer exists as  $s$  is  $Q$ -special.

We show that there is no subpath  $Q'$  of  $Q$  from  $v_i$  to some vertex  $v \notin V(G_\ell(s)) \setminus \text{pred}(s)$ . Observe that this implies the claimed statement as it shows that  $v_i \dots v_q$  does not leave  $G_\ell - \text{pred}(s)$ , and also that the same is true of  $v_i \dots v_1$ . Towards a contradiction suppose that such a path exists. Then by Lemma 5.2,  $Q'$  intersects  $X_s$ . Note that  $Q'$  does not intersect  $K^s$  if it is a subpath of  $v_i \dots v_q$  as otherwise there would be an edge between  $v_1$  and an internal vertex of  $Q'$ , whereas  $Q$  is induced. In the case where  $Q'$  is a subpath of  $v_i \dots v_1$ , it may intersect  $K^s$  on at most two vertices, possibly  $v_1$  and  $v_2$  where it stops. Thus  $Q'$  cannot leave  $G_\ell - \text{pred}(s)$  through the clique  $K^s$ . So it intersects the bag of a leaf  $t$  of  $B_\ell(s)$  and, in order to reach  $v_i$ , follows one of  $L(\pi(v_i), t)$  or  $R(\pi(v_i), t)$  (as defined in Section 3 when constructing the blow-up). Recall that  $v_1$  has neighbors in both such paths (these neighbors are endpoints of ribs incident to  $v_1$ ). This contradicts the fact that  $Q$  is induced.  $\square$

**Lemma 6.6.** *Let  $Q$  be an induced path of  $G_\ell$  and let  $a \in \{1, \dots, \ell\}$ . At most two sources of rank  $a$  in  $B_\ell$  are  $Q$ -special.*

*Proof.* Let  $Q = v_1 \dots v_q$ . We may assume that  $B_\ell$  has at least two  $Q$ -special sources of rank  $a$  as otherwise the statement holds. So there exist  $1 \leq i < j \leq \ell$  and two  $Q$ -special sources of rank  $a$  that we call  $s_i$  and  $s_j$  and such that  $v_i$  and  $v_j$  are respectively incident to the top edges of  $K^{s_i}$  and  $K^{s_j}$ . We choose them with  $j - i$  minimum. As  $s_i$  is a  $Q$ -special source, by Lemma 6.5, one of the following holds:

$$\{v_1, \dots, v_i\} \subseteq V(G_\ell(s_i)) \quad \text{or} \quad \{v_i, \dots, v_q\} \subseteq V(G_\ell(s_i)).$$

The same argument applied to  $s_j$  yields that one of the following holds:

$$\{v_1, \dots, v_j\} \subseteq V(G_\ell(s_j)) \quad \text{or} \quad \{v_j, \dots, v_q\} \subseteq V(G_\ell(s_j)).$$

However, the trees  $B_\ell(s_i)$  and  $B_\ell(s_j)$  only contain one source of rank  $a$ , respectively  $s_i$  and  $s_j$ . Therefore among the choices above only the two following outcome are possible:

$$\{v_1, \dots, v_i\} \subseteq V(G_\ell(s_i)) \quad \text{and} \quad \{v_j, \dots, v_q\} \subseteq V(G_\ell(s_j)).$$

So  $v_1 \dots, v_i$  and  $v_j \dots v_q$  contribute to two  $Q$ -special sources of rank  $a$  in  $Q$ . By minimality of  $j - i$ , the section of  $Q$  between  $v_i$  and  $v_j$  does not contribute to any  $Q$ -special source of rank  $a$ . Hence at most two sources of rank  $a$  of  $B_\ell$  are  $Q$ -special, as claimed.  $\square$

We are now ready to prove our main result, that we restate here.

**Theorem 1.7.** *There is a constant  $c$  such that for infinitely many integers  $n$ , there is a 2-degenerate graph  $G$  that has a path of order  $n$  and no induced path of order more than  $c(\log \log n)^2$ .*

*Proof.* We fix  $\ell \in \mathbb{N}$  and consider the graph  $G_\ell$ . Let  $q$  denote the maximum order of an induced path in  $G_\ell$ . We first show the following.

**Claim 6.7.** *There is a constant  $c$  such that  $q \leq c\ell^2$ .*

*Proof.* Let  $Q$  be an induced path of  $G_\ell$ . By Lemma 6.6, for every  $a \in \{1, \dots, \ell\}$  there are at most two sources of  $B_\ell$  of rank  $a$  that are  $Q$ -special. Therefore in total at most  $2\ell$  sources of  $B_\ell$  are  $Q$ -special. Let  $A \subseteq V(Q)$  be the set of vertices of  $Q$  such that  $\pi(v)$  is  $Q$ -special. For any node  $s \in V(B_\ell)$ ,  $|\pi^{-1}(s)| \leq 7$  (see Remark 3.2) hence  $|A| \leq 14\ell$ . We now consider a maximal subpath  $Q'$  of  $Q$  that does not intersect  $A$ . By definition of  $A$ , no source of  $B_\ell$  is  $Q'$ -special. Therefore by Lemma 6.1,  $|Q'| \leq c_{6.1} \cdot \ell$ . We conclude that  $|Q| \leq |A| + (2\ell + 1) \cdot c_{6.1} \cdot \ell \leq c \cdot \ell^2$  for some appropriately chosen constant  $c$ . Hence  $q \leq c\ell^2$ .  $\lrcorner$

Let  $n = |G_\ell|$ . As proved in Lemmas 4.1 and 4.3,  $G_\ell$  has a path of order  $n$  and has degeneracy 2. Also, by Lemma 4.2,  $n \geq 2^{2^{\ell+1}}$ . So  $c(\log \log n)^2 \geq c(\ell + 1)^2 \geq q$  (by Claim 6.7) proving the theorem.  $\square$

## 7 Bounding the coloring numbers of $G_\ell$

Given a total ordering  $\sigma$  of the vertices of a graph  $G$  and an integer  $r \geq 1$ , we say that a vertex  $y$  is  $r$ -reachable from a vertex  $x$  (with respect to  $\sigma$ ) if  $y <_\sigma x$  and there is a path  $P$  of length (i.e., number of edges) at most  $r$  such that every internal vertex  $z$  of  $P$  satisfies  $x < z$ .

The  $r$ -coloring number  $\text{col}_r(G)$  is the minimum integer  $k$  such that there is a total ordering  $\sigma$  on  $V(G)$  such that at most  $k$  vertices are  $r$ -reachable from every vertex of  $G$  [KY03].

The value  $\text{col}_1(G)$  is exactly the degeneracy of  $G$ . We saw in Lemma 4.3 that the  $G_\ell$ 's have degeneracy 2. Below we show that furthermore their coloring numbers are bounded from above by a linear function of  $r$ , i.e., Theorem 1.8, that we restate here for convenience.

**Theorem 1.8.**  *$\mathcal{G}$  has coloring numbers bounded from above by a linear function.*

*Proof.* Let  $\ell \in \mathbb{N}_{\geq 1}$ . We will prove that for every integer  $r \geq 1$ ,  $\text{col}_r(G_\ell) \leq 2r + 8$ . Let  $\sigma$  be any total order on  $V(G_\ell)$  such that for every  $u, v \in V(G_\ell)$ , if  $\pi(u) \prec \pi(v)$  then  $u <_\sigma v$ . Let  $r \in \mathbb{N}_{\geq 1}$ .

**Claim 7.1.** *Let  $x_1 \in V(G_\ell)$  and let  $x_1 \dots x_p$  (for some  $p \in \{2, \dots, r+1\}$ ) be a path witnessing that  $x_p$  is  $r$ -reachable from  $x_1$ . Then  $\pi(x_p) \preceq \pi(x_1)$  and for every  $i \in \{2, \dots, p-2\}$ ,  $\pi(x_1) \preceq \pi(x_i)$ .*

*Proof.* We start with the first part of the statement. Towards a contradiction, let us suppose that  $\pi(x_p) \not\preceq \pi(x_1)$ . By definition of reachability we have  $x_p <_\sigma x_1$ , so by definition of  $\sigma$  we deduce that  $\pi(x_1)$  and  $\pi(x_p)$  are  $\prec$ -incomparable. Recall that in  $G_\ell$ , if  $uv$  is an edge then  $\pi(u)$  and  $\pi(v)$  are  $\prec$ -comparable. So in  $G_\ell$ ,  $x_1$  and  $x_p$  are separated by the union of the bags of the common ancestors of  $\pi(x_1)$  and  $\pi(x_p)$ . Hence for some  $i \in \{2, \dots, p-1\}$ , the node  $\pi(x_i)$  is a common ancestor of  $\pi(x_1)$  and  $\pi(x_p)$ . But then  $x_i <_\sigma x_p$ , a contradiction. Hence  $\pi(x_p) \preceq \pi(x_1)$ . The argument is similar for the second part of the statement: assuming it does not hold for some  $i \in \{2, \dots, p-1\}$ , we deduce that  $\pi(x_1)$  and  $\pi(x_i)$  are  $\prec$ -incomparable and thus there is a  $j \in \{2, \dots, i-1\}$  such that  $\pi(x_j)$  is a common ancestor of  $\pi(x_1)$  and  $\pi(x_i)$ , implying the contradictory  $x_i <_\sigma x_1$ .  $\lrcorner$

Let  $x$  in  $V(G_\ell)$  and let  $d$  be the depth of  $x$ .

**Claim 7.2.** *At most 8 vertices  $y$  are  $r$ -reachable from  $x$  via a path whose edge incident to  $y$  is not a rib.*

*Proof.* Let  $y$  be such a vertex and let  $z$  be the other endpoint of the aforementioned edge. By definition we have  $y <_\sigma x \leq_\sigma z$  so by Claim 7.1 we get  $\pi(y) \preceq \pi(x) \preceq \pi(z)$ . Observe that if we ignore ribs, the graph in hand is merely the blow-up of a tree so the neighborhood of  $z$  is included in the union of the bag of  $\pi(z)$  and the bags of the nodes that are neighbors of  $\pi(z)$ . So the above inequality implies that  $y$  is either in the bag of  $\pi(x)$  or in the bag of the parent  $t$  of  $\pi(x)$ . We can further reduce the choices for  $y$  by observing that only two vertices of the bag of  $t$  have neighbors in the bag of  $\pi(x)$  and also by noting that  $y \neq x$ . This leaves at most  $7 + 2 - 1 = 8$  possible choices for  $y$ , as claimed.  $\lrcorner$

Let  $Y$  denote the set containing every vertex  $y$  that is  $r$ -reachable from  $x$  via a path whose edge incident to  $y$  is a rib. We call  $z_y$  the other endpoint of such an edge. So now in order to conclude the proof it is enough to prove the following statement.

**Claim 7.3.**  $|Y| \leq 2r$ .

*Proof.* Recall that if  $yz$  is a rib for some vertex  $z$  then  $(\text{depth}(y), \text{depth}(z)) \in \mathcal{N}_\ell$ . Let

$$(i_1, i'_1), \dots, (i_k, i'_k) \in \mathcal{N}_\ell$$

denote the intervals of the form  $(\text{depth}(y), \text{depth}(z_y))$  for  $y \in Y$ , ordered so that  $i_1 > \dots > i_k$ . By the property that intervals of  $\mathcal{N}_\ell$  do not cross (Remark 3.1) and the fact that  $\pi(x) \preceq \pi(z_y)$  for all such  $z_y$  (Claim 7.1), we have the following ordering:

$$i_k < \dots < i_1 \leq d \leq i'_1 < \dots < i'_k.$$

For the same reason there is no  $(i, i') \in \mathcal{N}_\ell$  with  $d \leq i < i'_j < i'$  for some  $j \in \{1, \dots, k\}$ , so no rib of  $G_\ell$  joins a vertex of depth  $i$  to one of depth  $i'$  for such  $i, i'$ . Therefore any path  $P$  with vertices in  $\{z \in V(G) : x \preceq_\sigma z\}$  from  $x$  to a vertex of depth  $i'_k$  contains a vertex of depth  $i'_j$  for every  $j \in \{1, \dots, k\}$ . This shows that  $P$  has length at least  $k - 1$ , hence  $k \leq r$ .

Recall that by Claim 7.1, every  $y \in Y$  belongs to the bag of an ancestor of  $\pi(x)$ . So  $Y$  is included in the union of the bags of the ancestors of  $\pi(x)$  of respective depths  $i_1, \dots, i_k$ . Furthermore, for each node  $t \in V(B_\ell)$ , only the two endpoints of the root edge of the bag of  $t$  are incident to ribs to vertices in the bags of deeper nodes. So we get  $|Y| \leq 2k \leq 2r$ , as claimed.  $\lrcorner$

From Claims 7.2 and 7.3 we get that  $\text{col}_r(G_\ell) \leq 2r + 8$ , as desired.  $\square$

## 8 Discussion

In this paper we described a construction disproving Conjecture 1.5 by tightening the upper-bound on the function  $f_k$  of Question 1.3 as follows:

$$\forall n, k \in \mathbb{N}_{\geq 2}, \quad \frac{\log \log n}{\log(k+1)} \leq f_k(n) \leq c(\log \log n)^2,$$

for some constant  $c > 0$ . So far there is no evidence that the  $O((\log \log n)^2)$  bound could be asymptotically tight so we did not try to optimize  $c$  in our proofs.<sup>8</sup> We however note that our bound cannot be asymptotically improved by conducting a different analysis on the same construction, as our graphs contain induced paths of order  $\Omega((\log \log n)^2)$ ; see Figure 6 for an example. The main open problem following this work is to determine the exact order of magnitude of  $f_k$ . An other (possibly more difficult) problem is to find the optimal bound in Theorem 1.2.

---

<sup>8</sup>It seems that the same construction starting with a  $k$ -ary tree instead of a binary one, we could win a  $\frac{1}{\log(k+1)}$  factor for  $k$ -degenerate graphs.

Let us now ask a more general question. For a graph class  $\mathcal{G}$  that excludes some biclique as subgraph,<sup>9</sup> let  $f_{\mathcal{G}}: \mathbb{N} \rightarrow \mathbb{R}$  be the maximum function such that for every  $G \in \mathcal{G}$ , if  $G$  has a path of order  $n$  then it has an induced path of order at least  $f_{\mathcal{G}}(n)$ . In this paper we proved that there are classes  $\mathcal{G}$  with linearly bounded coloring numbers (hence linearly bounded expansion) and where  $f_{\mathcal{G}}(n) = O((\log \log n)^2)$ . On the other hand, for several graph classes  $\mathcal{G}$  it is known that  $f_{\mathcal{G}}(n) = \Omega((\log n)^c)$  for some  $c > 0$ . Such bounds exist for instance for planar graphs (Theorem 1.4), graph classes of bounded treewidth, more generally any class that excludes a topological minor (Theorem 1.6). This exhibits two fundamentally different behaviors for the function  $f_{\mathcal{G}}$ . In [HR23] a third behavior is identified: graph classes where  $f_{\mathcal{G}}$  is polynomial, such as graphs of bounded pathwidth. It is therefore natural to ask: which criteria distinguish graph classes  $\mathcal{G}$  such that  $f_{\mathcal{G}}$  is polynomial, polylogarithmic, and doubly-polylogarithmic?

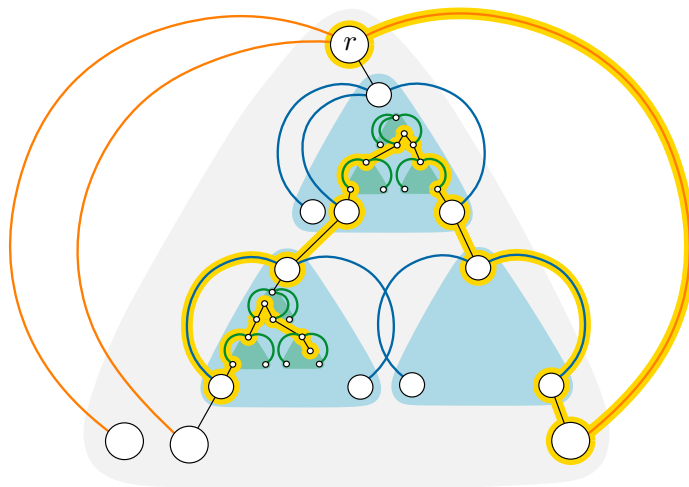


Figure 6: A representation (highlighted in gold) of an induced path of order  $\Omega((\log \log n)^2)$  in  $G_\ell$  starting from its root bag. For readability, only its "trace" in  $B_\ell$  is represented together with the intervals (here drawn in a curved way) it follows, i.e., the actual path in  $G_\ell$  may be obtained by replacing each highlighted node  $s \in V(B_\ell)$  by appropriate vertices in  $K^s$ , each highlighted edge  $ss' \in E(B_\ell)$  by an appropriate path  $L(s, s')$  or  $R(s, s')$ , and each highlighted interval by an appropriate rib it originates from. A careful analysis shows that the order of such a path is  $\Omega(\ell^2)$ .

<sup>9</sup>As discussed in the introduction, excluding arbitrarily large cliques and bicliques is necessary (in hereditary classes) for such a function to exist.

## Acknowledgements

We thank Louis Esperet for suggesting us to investigate whether our construction has bounded expansion, which led to Section 7.

## References

- [ELM17] Louis Esperet, Laetitia Lemoine, and Frédéric Maffray. Long induced paths in graphs. *European Journal of Combinatorics*, 62:1–14, 2017.
- [ER18] Louis Esperet and Jean-Florent Raymond. Polynomial expansion and sublinear separators. *European Journal of Combinatorics*, 69:49 – 53, 2018.
- [GRS82] Fred Galvin, Ivan Rival, and Bill Sands. A Ramsey-type theorem for traceable graphs. *Journal of Combinatorial Theory, Series B*, 33(1):7–16, 1982.
- [HR23] Claire Hilaire and Jean-Florent Raymond. Long induced paths in minor-closed graph classes and beyond. *Electronic Journal of Combinatorics*, 63(1), 2023.
- [KY03] Hal A. Kierstead and Daqing Yang. Orderings on graphs and game coloring number. *Order*, 20:255–264, 2003.
- [NOdM12] Jaroslav Nešetřil and Patrice Ossona de Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer Science & Business Media, 2012.