

Long induced paths in minor-closed graph classes and beyond

Claire Hilaire

Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800
33400, Talence, France

`claire.hilaire@u-bordeaux.fr`

Jean-Florent Raymond*

CNRS, LIMOS, Université Clermont-Auvergne
63000 Clermont-Ferrand, France

`jean-florent.raymond@cnrs.fr`

Submitted: Feb 4, 2022; Accepted: Nov 30, 2022; Published: Jan 27, 2023

© The authors. Released under the CC BY license (International 4.0).

Abstract

In this paper we show that every graph of pathwidth less than k that has a path of order n also has an induced path of order at least $\frac{1}{3}n^{1/k}$. This is an exponential improvement and a generalization of the polylogarithmic bounds obtained by Esperet, Lemoine and Maffray (2016) for interval graphs of bounded clique number. We complement this result with an upper-bound.

This result is then used to prove the two following generalizations:

- every graph of treewidth less than k that has a path of order n contains an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$;
- for every non-trivial graph class that is closed under topological minors there is a constant $d \in (0, 1)$ such that every graph from this class that has a path of order n contains an induced path of order at least $(\log n)^d$.

We also describe consequences of these results beyond graph classes that are closed under topological minors.

Mathematics Subject Classifications: 05C38

*This author was supported by the ANR project GRALMECO (ANR-21-CE48-0004).

1 Introduction

In this paper we are concerned with the dependency between the maximum lengths of paths and induced paths in graphs. Every induced path is a path. Conversely, does the existence of a long path in a graph imply that of a long induced path? This question may be formalized as asking for the existence of an increasing function f such that the following property holds:

- ($\star$) *For every $n \in \mathbb{N}$, if a graph has a path of order n , then it has an induced path of order at least $f(n)$.*

In general, such a function does not exist, as shown by cliques and bicliques. In 1982, Galvin, Rival, and Sands showed that these are the only obstructions by providing an increasing function satisfying ($\star$) for graphs excluding a biclique as a subgraph [GRS82, Theorem 4]. Their proof relies on the infinite Ramsey's theorem for 4-tuples, hence their result should first and foremost be seen as an existential result, rather than one providing accurate and tight bounds.

A better lower-bound was given in [NDM12, Lemma 6.4] by Nešetřil and Ossona de Mendez in the case of k -degenerate¹ graphs: there, a path of order n implies the existence of an induced path of order at least $\frac{\log \log n}{\log(k+1)}$. They also proposed as an open problem to find the maximum function $f: \mathbb{N}^2 \rightarrow \mathbb{N}$ such that, for every $k \in \mathbb{N}$, property ($\star$) holds for k -degenerate graphs with bound $n \mapsto f(k, n)$ [NDM12, Problem 6.1].

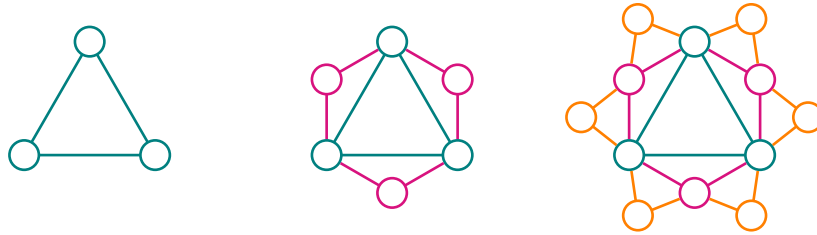


Figure 1: First terms of a sequence of outerplanar graphs where induced paths have length at most logarithmic in the order of the graph.

A graph is said to be *outerplanar* if it can be drawn in the plane with no edge crossing and with all the vertices on the outer face. Outerplanar graphs form a subclass of planar graphs and are known to be 2-degenerate. Already within this simple class there are infinite families of graphs where the order of the longest induced path is at most logarithmic in the order of the graph [AV00], see Figure 1 for an example. This motivates the research for classes of degenerate graphs where a polylogarithmic² lower-bound for the function of property ($\star$) could be obtained. As a possible answer to the aforementioned question of Nešetřil and Ossona de Mendez, it was conjectured in [ELM17] that this would be the *correct* bound for k -degenerate graphs.

¹For a positive integer k , a graph G is said to be k -degenerate if every subgraph of G (including G itself) has a vertex of degree at most k .

²That is, of the form $n \mapsto c(\log n)^d$ for some reals $c, d > 0$.

Conjecture 1 ([ELM17, Conjecture 1.1]). For every integer k there is a constant d such that every k -degenerate graph that has a path of order n also has an induced path of order at least $(\log n)^d$.

In 2000, Arocha and Valencia [AV00] and later Di Giacomo, Liotta and Mchedlidze [DLM16] considered 3-connected planar graphs and 2-connected outerplanar graphs.

Their results have then been improved by Esperet, Lemoine, and Maffray who showed that 2-connected outerplanar graphs and 3-connected planar graphs of order n have an induced path of order $\Omega(\log n)$ [ELM17, Theorems 3.2 and 3.5].³ They used these results to prove the following statement.

Theorem 2 ([ELM17, Theorem 3.8]). *For every $g \in \mathbb{N}$ there is a constant $c = \frac{1}{2\sqrt{6}} - o(1)$ such that for every graph G embeddable in a surface with Euler genus at most g , if G has a path of order n , then G has an induced path of order at least $c\sqrt{\log n}$.*

Besides planar graphs, a prominent type of degenerate graphs consists of graphs of bounded treewidth (to be defined in Section 2). Indeed, every graph of treewidth at most k is k -degenerate. If G has treewidth k and any addition of an edge between two vertices of G yields a graph of treewidth larger than k , then G is called a k -tree. In the same paper, Esperet et al. obtained the following bound for k -trees.

Theorem 3 ([ELM17, Theorem 2.2]). *For every k , if a k -tree has a path of order n then it has an induced path of order at least $\frac{\log n}{k \log k}$.*

Using similar ideas they were able to show the following.

Theorem 4 ([ELM17, Corollary 2.5]). *If a graph G of treewidth at most 2 has a path of order n then it has an induced path of order at least $(\frac{1}{2} - o(1)) \log n$.*

However it is not clear for larger values of k how the ideas of the proof of Theorem 3 could be adapted to the more general setting of graphs of treewidth less than k . Actually the upper-bound of $(k+1)(\log n)^{\frac{2}{k-1}}$ on the function of property $(\star)$ for graphs of treewidth less than k provided in the same paper even suggests that new techniques will be needed to approach such graphs.

The second type of class of graphs of bounded treewidth considered in the same paper is interval graphs of bounded clique number.⁴

Theorem 5 ([ELM17, Theorem 4.1]). *For every integer k there is a constant c such that, if G is an interval graph with clique number at most k and that has a path of order n , then G has an induced path of order at least $c(\log n)^{\frac{1}{(k-1)^2}}$.*

³These results do not require a long path to exist in the considered graph because this follows from the connectivity requirement in the considered graph classes, see e.g. [CY02].

⁴The *clique number* of a graph is the maximum order of a clique it contains. An interval graph of clique number k has treewidth less than k .

1.1 Our contribution

Our first result is a generalization of Theorem 5 to graphs of bounded pathwidth⁵, with an exponential improvement of the bound.

Theorem 6. *For every $k \in \mathbb{N}$, if G is a graph of pathwidth less than k that has a path of order n , then G has an induced path of order at least $\frac{1}{3}n^{1/k}$.*

Its proof consists in a simple win/win strategy. We identify an induced path whose removal decreases the pathwidth of G . Then either this path is at least as long as the bound promised by the statement and we are done, or its removal decreases the pathwidth without decreasing the number of vertices much, and we conclude by applying the induction hypothesis. Theorem 6 is complemented by an upper-bound (that even holds for interval graphs) of $n^{2/k} + 1$ (Theorem 21), showing that the exponential dependency in $1/k$ in our lower-bound above is unavoidable.

We then show that in a graph of small treewidth that has a large path there is always (as a contraction) a graph of small pathwidth that has a long path. This statement is used to obtain the following polylogarithmic bound for graphs of bounded treewidth.

Theorem 7. *For every $k \in \mathbb{N}$, if G is a graph of treewidth less than k that has a path of order n , then G has an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$.*

As mentioned above, Esperet et al. [ELM17] constructed chordal graphs of clique number k that have a path of order n and where no induced path has order more than $(k+1)(\log n)^{\frac{2}{k-1}}$. Therefore neither the logarithmic dependency in n nor the exponential dependency in $1/k$ could be improved in our lower-bound above.

The operation of *subdividing* an edge uv of a graph consists in deleting the edge and adding a new vertex of degree 2 adjacent to u and v . A *subdivision* of a graph H is a graph obtained by applying a (possibly empty) sequence of edge subdivisions to H . If a graph G has a subgraph that is (isomorphic to) a subdivision of H , we say that H is a *topological minor* of G . A graph class is said to be *closed under taking topological minors* if every topological minor of a graph of the class also belongs to the class. It is *non-trivial* if it is not the class of all graphs.

Non-trivial graph classes that are closed under topological minors form a very general setting that encompasses several fundamental graph classes such as any minor-closed graph class (e.g. planar graphs or more generally graphs of bounded Euler genus, and graphs of bounded pathwidth or treewidth), any immersion-closed graph class (like graphs of bounded cutwidth, carving width, or tree-cut width), and graphs of bounded degree.

The ideas developed in the proofs of Theorem 6 and Theorem 7 allowed us to generalize their statement to this much more general setting.

Theorem 8. *For every non-trivial graph class $\mathcal{G}$ that is closed under taking topological minors there is a constant $d \in (0, 1)$ such that if a graph $G \in \mathcal{G}$ has a path of order n , then G has an induced path of order at least $(\log n)^d$.*

⁵An interval graph has clique number k iff it has pathwidth $k - 1$.

The proof deals separately with the different parts that compose graphs excluding a topological minor (as given by the structure theorem of Grohe and Marx [GM15]) and then shows how they can be combined together. We actually prove a stronger statement than Theorem 8, namely that the same outcome holds for all graphs that admit a tree decomposition where every torso is either almost embeddable in a surface of bounded genus or has almost bounded degree (see Theorem 37 for the formal statement).

Before proceeding to the proofs, let us first mention some consequences of our results beyond graph classes that are closed under topological minors.

1.2 Consequences of our results

Graphs of bounded chordality. Gartland, Lokshtanov, Pilipczuk, Pilipczuk, and Rzażewski proved in [GLP⁺21] that in graphs of bounded chordality (i.e. graphs with no long induced cycles), degeneracy and treewidth are tied in the following sense.

Theorem 9 ([GLP⁺21, Theorem 1.3]). *For every $k, \ell \in \mathbb{N}$ there is a constant t such that if a graph G has degeneracy at most k and no induced cycle on at least ℓ vertices then G has treewidth at most t .*

Together with the above result, Theorem 7 implies Conjecture 1 for graphs of bounded chordality.

Corollary 10. *For every $k, \ell \in \mathbb{N}$ there is a constant $d \in (0, 1)$ such that if a graph G is k -degenerate with no induced cycle of order at least ℓ and G has a path of order n , then G has an induced path of order at least $(\log n)^d$.*

Graphs of bounded cliquewidth. Cliquewidth is a graph parameter that is more general than treewidth.⁶ However, if we forbid arbitrarily large cliques and bicliques then each of the two parameters can be upper-bounded by a function of the other [CO00]. Specifically, any $K_{s,s}$ -subgraph free graph of cliquewidth less than k has treewidth less than $3(k-1)(s-1)$, as proved in [GW00, Corollary 1]. Thanks to this property we have the following consequence of Theorem 7.

Corollary 11. *Let $k, s \in \mathbb{N}^+$, if G is a $K_{s,s}$ -subgraph free graph of cliquewidth less than k that has a path of order n , then G has an induced path of order at least $\frac{1}{4}(\log n)^{\frac{1}{3(k-1)(s-1)}}$.*

Clearly, forbidding $K_{s,s}$ -subgraphs (which amounts to forbidding large induced bicliques and large cliques) is a necessary condition in the statement of Corollary 11: as we mentioned above, in cliques and bicliques the property $(\star)$ only holds with a constant bound.

⁶Cliquewidth is more general in the sense that every class of graphs of bounded treewidth has bounded cliquewidth, while the converse does not hold in general. Cliquewidth only appears in the statement of Corollary 11 so we refrain from giving its quite lengthy definition here and instead refer the interested reader to one of the papers on the topic such as [CO00].

Well-quasi-ordered induced subgraphs ideals. Let $\mathcal{G}$ be a graph class. We say that $\mathcal{G}$ is *hereditary* if any induced subgraph of a graph of $\mathcal{G}$ also belongs to $\mathcal{G}$. The class $\mathcal{G}$ is *well-quasi-ordered by induced subgraphs* if in every infinite sequence $(G_i)_{i \in \mathbb{N}}$ of graphs from $\mathcal{G}$, there are integers $i < j$ such that G_i is an induced subgraph of G_j . Theorem 7 also has the following consequence.

Corollary 12. *Let $\mathcal{G}$ be a hereditary class of graphs that does neither contain all cliques nor all bicliques. If $\mathcal{G}$ is well-quasi-ordered by induced subgraphs then there is a constant $d \in (0, 1)$ such that the following holds: every $G \in \mathcal{G}$ that has a path of order n also has an induced path of order at least n^d .*

This holds because, as proved in [ALR19, Theorem 1], graph classes such as $\mathcal{G}$ in the above statement have pathwidth bounded from above by a constant (that depends on the order of the smallest excluded clique and biclique), so we can apply Theorem 7.⁷ Similarly as above the requirement in Corollary 12 that $\mathcal{G}$ does not contain all cliques or bicliques is necessary for hereditary classes. Note nevertheless that the assumption that $\mathcal{G}$ is well-quasi-ordered is only sufficient. For instance, the class of graphs of maximum degree at most 2 (which are unions of cycles and paths) is hereditary, excludes large cliques and bicliques, and satisfies $(\star)$ with a linear bound, however it is not well-quasi-ordered by induced subgraphs due to the presence of arbitrarily long cycles.

Graphs with polynomial-sized modulators. As we will prove in Section 5.1, graphs that are not *too far* from a class where $(\star)$ holds with a polylogarithmic bound also satisfy this property with a similar bound. More formally, we will prove the following statement.

Lemma 13 (Restatement of Lemma 27). *Let $\mathcal{G}$ be a hereditary class of graphs where $(\star)$ holds with function $n \mapsto c(\log n)^d$. Let $\varepsilon \in (0, 1)$ and let $\mathcal{G}_\varepsilon$ denote the class of graphs G that have a set of vertices X with $|X| \leq |G|^\varepsilon$ such that $G - X \in \mathcal{G}$.*

Then $\mathcal{G}_\varepsilon$ satisfies $(\star)$ with function $n \mapsto c'(\log n)^d$, for some constant c' that depends on c , d , and ε only.

Therefore all our results can be extended to classes that are, informally, at *polynomial distance* from the considered classes. As an example, the lemma above and Theorem 7 together imply the following result.

Corollary 14. *There is a constant $c \in (0, 1)$ such that the following holds. Let $k \in \mathbb{N}$ and let $\mathcal{G}_k$ denote the class of graphs G that have a set $X \subseteq V(G)$ with $|X| \leq \sqrt{|G|}$ such that $\text{tw}(G - X) \leq k$. Then for every $G \in \mathcal{G}_k$, if G has a path of order n then it has an induced path of order at least $c(\log n)^{1/k}$.*

Observe that $\mathcal{G}_k$ above is not closed under topological minors, so it does not fall under the umbrella of Theorem 8. Also, a direct application of Theorem 7 to graphs of $\mathcal{G}_k$ (with bound $k + \sqrt{|G|}$ on the treewidth) would not guarantee more than constant-sized induced paths.

⁷The bound on the pathwidth appeared in the published version of [ALR19]. As of the time of writing, the preprint of the paper on arxiv only proves a bound on the treewidth.

Graphs of moderate degree. In order to prove Theorem 8 we will show that graphs of degree bounded by a constant satisfy property $(\star)$ with a polylogarithmic bound (Corollary 30). More generally, we can prove that the same holds for graphs with subpolynomial maximum degree (Lemma 29). For instance, an application of this lemma shows that the class of the graphs G where the degree is at most $2\sqrt{\log |G|}$ satisfies $(\star)$ with bound $n \mapsto \sqrt{\log n}$. Note that this graph class is not closed under taking topological minors.

1.3 Organization of the paper.

In Section 2 we give the necessary definitions. We prove Theorem 6 in Section 3 and Theorem 7 in Section 4. Section 5 is split into several subsections in order to finally prove Theorem 8. We discuss further research directions in Section 6.

2 Preliminaries

Unless otherwise specified, logarithms are binary. Graphs in this paper are finite, undirected, and loopless. For every graph G we respectively denote by $V(G)$ and $E(G)$ its sets of vertices and edges and use $|G|$ as a shorthand for $|V(G)|$, the *order* of the graph. We use standard graph theory terminology.

Remark 15. If G is a graph that has a path P of order n and $X \subseteq (G)$ is not empty, then $G - X$ has

1. at most $|X| + 1$ connected components that contain a vertex of P ; and
2. a connected component that has a path of order at least $\frac{n-|X|}{|X|+1} \geq \frac{n}{2|X|} - 1$.

Long paths versus Hamiltonian paths. A *Hamiltonian path* in a graph G is a path that visits all the vertices of G . The statements of the results in Section 1 follow the general form of $(\star)$. In contrast, in the rest of the paper we work with statements of the following form (for $\mathcal{G}$ a class of graphs):

$(\star\star)$ *For every $n \in \mathbb{N}$, if a graph of $\mathcal{G}$ has a Hamiltonian path and order n then it has an induced path of order at least $f(n)$.*

Remark 16. For a hereditary graph class $\mathcal{G}$, the statements $(\star)$ and $(\star\star)$ are equivalent. Given a graph $G \in \mathcal{G}$ with a path P of order n (as required by $(\star)$), it suffices to consider the induced subgraph $G[V(P)] \in \mathcal{G}$ to be able to apply statement $(\star\star)$. This shows that $(\star\star)$ implies $(\star)$ and the other direction is trivial.

However the form $(\star\star)$ is more convenient for the proofs because the induced paths that we construct will never use vertices other than those of the path P whose existence is assumed, so by using form $(\star\star)$ we do not need to explicitly say that we restrict our attention to $G[V(P)]$.

Representations. To easily deal with graphs of bounded pathwidth or treewidth (defined hereafter) we find it convenient to define tree representations, which are objects that are closely related to tree decompositions, as we explain below. Formally, a *tree representation* of a graph G is a pair $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ such that

1. T is a tree;
2. for every $v \in V(G)$, T_v is a subtree of T , called the *model* of v ;
3. for every edge uv of G the subtrees T_u and T_v intersect.

When T is a path, we call $\mathcal{T}$ a *path representation*. If item (3) is strengthened as follows

- 3'. uv is an edge of G if and only if T_u and T_v intersect

then G is the intersection graph of the vertex sets of the subtrees $\{T_v\}_{v \in V(G)}$ and it is a *chordal graph*⁸; if furthermore T is a path then G is an *interval graph*. In this latter case we call $\mathcal{T}$ an *interval representation* of G to stress that it is an interval graph. To avoid confusion between the vertices of G and T , we use the synonym *nodes* to refer to vertices of T .

Tree representations and tree decompositions are closely linked, as we explain now. For a tree representation $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ of a graph G , we define for every $t \in V(T)$ the *bag* at t as the following subset of $V(G)$

$$\beta_{\mathcal{T}}(t) = \{v \in V(G) \mid t \in T_v\}.$$

We drop the subscript when there is no ambiguity. The *width* of a tree representation $\mathcal{T}$ is $\max_{t \in V(T)} |\beta_{\mathcal{T}}(t)| - 1$. The *treewidth* of G is the minimum width of any of its tree representations and the *pathwidth* of G is the minimum width of any of its path representations. It can easily be seen that these definitions coincide with the usual definitions for treewidth and pathwidth as, with the notation above, $(T, \{\beta(t)\}_{t \in V(T)})$ is a tree decomposition of G . We respectively denote by $\mathbf{tw}(G)$ and $\mathbf{pw}(G)$ the treewidth and pathwidth of G .

The following is a consequence of items (2) and (3) of the definition of a tree representation.

Remark 17. Let G be a graph with a tree representation $(T, \{T_v\}_{v \in V(G)})$ and let H be a connected subgraph of G . Then $\bigcup_{v \in V(H)} T_v$ is a (connected) subtree of T .

3 Induced paths in graphs of bounded pathwidth

In this section we show that the maximum function f_k such that property $(\star)$ holds for graphs of pathwidth less than k is such that $\frac{1}{3}n^{1/k} \leq f_k(n) \leq n^{2/k} + 1$ (Theorem 6 and Theorem 21).

⁸Chordal graphs are usually defined as graphs with no induced cycle of order 4 or more. They can be seen as intersection graphs of subtrees of a tree, as proved by Gavril [Gav74, Theorem 3].

A *caterpillar* is a tree in which all the vertices are at distance at most one of some path.

Lemma 18 (See [PT99, Section 6]). *Every connected graph of pathwidth at most one is a caterpillar.*

We will also use the following consequence of the Helly property of intervals (see for instance [GGL95, Section 2.5]).

Lemma 19. *Suppose $\mathcal{P} = (P, \{P_v\}_{v \in V(G)})$ is a path representation of a graph G and K is a clique of G . Then there is a node $t \in V(P)$ such that $V(K) \subseteq \beta_{\mathcal{P}}(t)$.*

We first prove Theorem 6, that we restate here for convenience.

Theorem 6. *For every $k \in \mathbb{N}$, if G is a graph of pathwidth less than k that has a path of order n , then G has an induced path of order at least $\frac{1}{3}n^{1/k}$.*

Proof. We actually prove the following by induction on n and k , which is equivalent to the desired statement according to Remark 16.

If G is a graph of order at least n and pathwidth less than k that has a Hamiltonian path, then G has an induced path of order at least $\frac{1}{3}n^{\frac{1}{k}}$.

Note that the statement is vacuously true when $k = 1$ and $n > 1$ as there is no connected graph with $|G| \geq 2$ and $\mathbf{pw}(G) = 0$. When $\mathbf{pw}(G) = n - 1$ (thus $k = n$), the graph is a clique and the order of its longest induced path is $2 \geq \frac{1}{3}n^{\frac{1}{k}}$. In the cases where $n \geq 2$ and $k = 2$, G is a caterpillar, by Lemma 18. Every caterpillar with a Hamiltonian path is a path, so the statement holds as G is already an induced path and $n \geq \frac{1}{3}\sqrt{n}$.

So we now assume that $k \geq 3$, $n > k$ and that the statement holds for all smaller values of k and n . We now show that it also true for k and n . We may also assume that $\frac{1}{3}n^{\frac{1}{k}} > 2$ as otherwise any edge is an induced path of the required order.

Let G be a graph that has a Hamiltonian path and such that $|G| \geq n$ and $\mathbf{pw}(G) < k$. We fix a path representation $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ of G of width less than k and with no empty bag. Let u be a vertex of G such that R_u contains one endpoint of R and let v be a vertex whose model contains the other endpoint; possibly $u = v$. Let Q be an induced path in G between u and v . Observe that Q may consist of a single vertex (when $u = v$) or of two vertices (when u and v are adjacent). Otherwise, Q can be constructed by taking a shortest path between u and v .

In the case where $|Q| \geq \frac{1}{3}n^{\frac{1}{k}}$ we are done and Q is the desired path. So in the rest of the proof we may assume that $|Q| < \frac{1}{3}n^{\frac{1}{k}}$. Let P be a Hamiltonian path of G . The path Q intersects P in $|Q|$ vertices thus removing Q from G cuts P into at most $|Q| + 1$ subpaths, and the longest of them, that we call P' , has an order n' that is at least $\frac{n - |Q|}{|Q| + 1}$. Let us consider G' the graph induced by P' in G .

We now show that $\mathbf{pw}(G') \leq k - 1$. For this we consider the path representation $\mathcal{R}' = (R, \{R_w\}_{w \in V(G')})$ of G' . Let $r \in R$ be a node such that $|\beta_{\mathcal{R}'}(r)|$ is maximum. From the definition we have $\beta_{\mathcal{R}'}(r) \subseteq \beta_{\mathcal{R}}(r)$. By Remark 17, the definition of Q and the fact

that it is connected, we know that the union of the sets $\{R_w\}_{w \in V(Q)}$ is equal to $V(R)$. Therefore $\beta_{\mathcal{R}}(r)$ contains a vertex of Q . This implies $|\beta_{\mathcal{R}}(r)| > |\beta_{\mathcal{R}'}(r)|$ and the claimed bound on the pathwidth of G' follows.

The graph G' has a Hamiltonian path (by definition), order $n' < n$ and pathwidth at most $k - 1$. By induction, G' admits an induced path Q' of order at least $\frac{1}{3}n'^{\frac{1}{k-1}}$. Recall that $n' \geq \frac{n-|Q|}{|Q|+1}$. Since $|Q| < \frac{1}{3}n^{\frac{1}{k}}$, we have $n' > \frac{n}{\frac{1}{3}n^{\frac{1}{k}}+1} - 1$. As we assume $\frac{1}{6}n^{\frac{1}{k}} \geq 1$ we deduce $n' > 2n^{\frac{k-1}{k}} - 1 \geq n^{\frac{k-1}{k}}$. Therefore we have $|Q'| \geq \frac{1}{3}(n^{\frac{k-1}{k}})^{\frac{1}{k-1}} = \frac{1}{3}n^{\frac{1}{k}}$. Since G' is an induced subgraph of G , Q' is also an induced path of G so we are done. $\square$

As every interval graph of clique number k admits a path representation of width less than k (given by its interval representation), we have the following improvement of the bound of Theorem 5.

Corollary 20. *For every $k, n \in \mathbb{N}$, if G is an interval graph of order at least n and clique number at most k that has a Hamiltonian path, then G has an induced path of order at least $\frac{1}{3}n^{\frac{1}{k}}$.*

The following statement complements Theorem 6 and Corollary 20 by giving an upper-bound on the order of induced paths one can guarantee in (interval) graphs of bounded pathwidth.

Theorem 21. *For every $k, n \in \mathbb{N}$ with $2 \leq k \leq n$, there exists an interval graph $G_{n,k}$ with a Hamiltonian path of order at least n and clique number at most k , such that every induced path of $G_{n,k}$ has order at most $n^{\frac{2}{k}} + 1$.*

Proof. The proof is by induction on n and k . When $k = n$, $G_{n,n}$ is the clique on n vertices, where every induced path has order at most 2, which is less than $n^{\frac{2}{n}} + 1$. When $k = 2$, $G_{n,2}$ is the (induced) path on n vertices, which again satisfies the desired statement for every $n \geq 2$.

For any $n \geq k$, let $q = \left\lfloor n^{\frac{2}{3}} + 1 \right\rfloor$. When $k = 3$, $G_{n,3}$ is constructed from a collection $\{P_i\}_{i \in \{1, \dots, q\}}$ of q paths, where for every $i \in \{1, \dots, q\}$ the path P_i has order i and an endpoint called u_i , by connecting u_i to all the vertices of P_{i+1} for every $i \in \{1, \dots, q-1\}$. See Figure 2b for a depiction of a small case and Figure 2a for an interval representation $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ of it, where each R_v is drawn below R respecting the x-axis.

We can see that at most 3 intervals intersect (R_{u_i} and the intervals representing two consecutive vertices of P_{i+1} for each i), thus $G_{n,3}$ is an interval graph with clique number 3.

The number of vertices of $G_{n,3}$ is $\sum_{i=1}^q |P_i| = \frac{q(q+1)}{2} \geq \frac{n^{\frac{2}{3}}(n^{\frac{2}{3}}+1)}{2}$ since $q \geq n^{\frac{2}{3}}$, which is greater than n . $G_{n,3}$ admits a Hamiltonian path starting in u_q that, for each i from q to 2, follows P_i from u_i to its other endpoint, goes to u_{i-1} , and repeats the same process.

Let us now bound the maximum order of an induced path in $G_{n,3}$. Let Q be an induced path of $G_{n,3}$ and let i denote the minimum integer such that Q has a vertex from

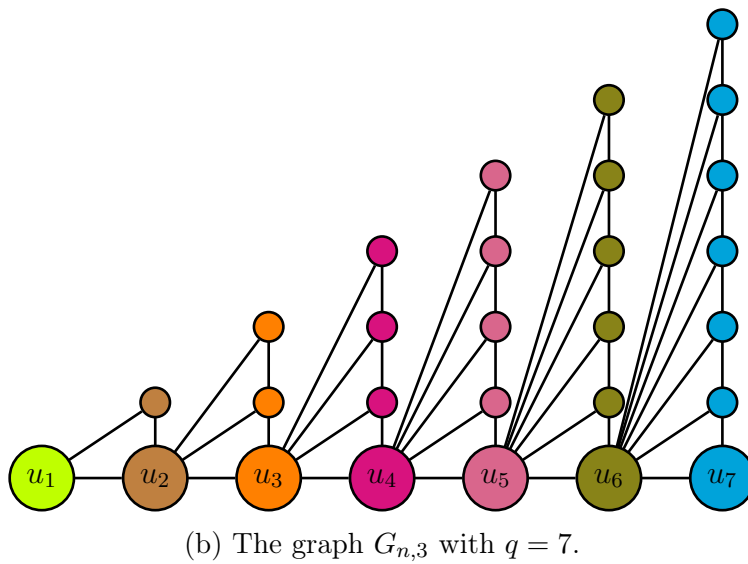
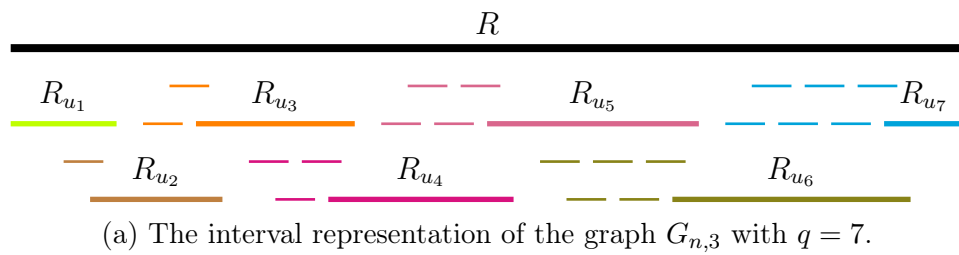


Figure 2: The construction of Theorem 21.

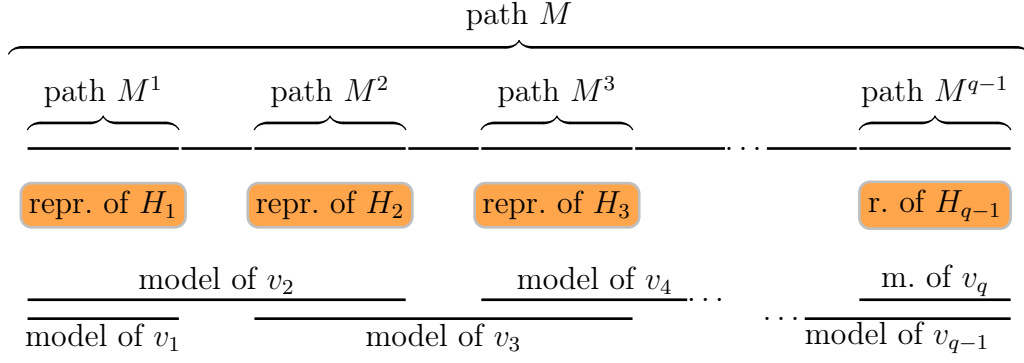


Figure 3: The construction of an interval representation for $G_{n,k}$ when $k > 3$.

P_i ; clearly Q has at most i vertices from this path. Let $j \in \{i+1, \dots, q\}$ and observe that every vertex of P_j has u_{j-1} as unique neighbor in $\{P_{j'}\}_{j' < j}$. As Q is induced we deduce that it contains at most one vertex of P_j . This holds for each of the $q-i$ paths of $\{P_j\}_{j \in \{i+1, \dots, q\}}$, so we get the bound $|Q| \leq q$, as desired. This concludes the proof for the case $k = 3$.

So we now take $n > k \geq 4$, and assume that $G_{n',k'}$ is defined and satisfies the statement for every $k' < k$ and $n' < n$ such that $2 \leq k' \leq n'$. To construct $G_{n,k}$, we proceed as follows.

Let $q = \left\lfloor n^{\frac{2}{k}} + 1 \right\rfloor$. If $q \geq n$ then the graph $G_{n,k} = P_q$ clearly satisfies the desired statement. Otherwise, we set $n' = \left\lceil \frac{n-q}{q-1} \right\rceil$; observe that $n' \geq 2$. We construct $G_{n,k}$ from the disjoint union of a path $Q = v_1 \dots v_q$ and $q-1$ copies $H_1, \dots, H_{q-1}$ of $G_{n',k-2}$ by connecting v_i and v_{i+1} to all vertices of H_i , for every $i \in \{1, \dots, q-1\}$.

To show that $G_{n,k}$ is an interval graph, we now provide an interval representation of it (see Figure 3 for an illustration). By our induction hypothesis, H_i admits an interval representation $\mathcal{M}^i = (M^i, \{M_v^i\}_{v \in V(H_i)})$ for every $i \in \{1, \dots, q-1\}$. Let M denote the concatenation of the paths $M^1, \dots, M^{q-1}$. That is, M is obtained from the disjoint union of these paths by adding an edge between an endpoint of M^1 and one endpoint of M^2 , from the other endpoint of M^2 to one endpoint of M^3 , and so on. Clearly $(M, \bigcup_{i=1}^{q-1} \{M_v^i\}_{v \in V(H_i)})$ is an interval representation of the disjoint union of the H_i 's. For every $i \in \{2, \dots, q-1\}$, let $M_{v_i} = M[V(M^{i-1}) \cup V(M^i)]$ and let $M_{v_1} = M^1$ and $M_{v_q} = M^{q-1}$. Observe that $(M, \{M_{v_i}\}_{i \in \{1, \dots, q\}})$ is an interval representation of Q . Furthermore, $\mathcal{M} = (M, \{M_{u_i}\}_{i \in \{1, \dots, q\}} \cup \bigcup_{i=1}^{q-1} \{M_v^i\}_{v \in V(H_i)})$ is an interval representation of $G_{n,k}$. This proves that $G_{n,k}$ is an interval graph.

The graph $G_{n,k}$ is composed of the path Q on q vertices and $q-1$ copies of $G_{n',k-2}$ which have (by induction) at least n' vertices each, so we have

$$\begin{aligned} |G_{n,k}| &\geq q + (q-1)n' \\ &= q + (q-1) \left\lceil \frac{n-q}{q-1} \right\rceil \\ &\geq n. \end{aligned}$$

We now show that the clique number of $G_{n,k}$ is at most k . Let K be a maximum clique of $G_{n,k}$. By Lemma 19, there exists a node x of M such that K consists of all vertices of G whose model (in $\mathcal{M}$) contains x . Let j be such that $x \in M^j$ (such a vertex exists by definition of M), we then have

$$V(K) = (V(K) \cap V(Q)) \cup (V(K) \cap V(H_j)).$$

The first intersection has size at most 2 since Q is an induced path. Recall that H_j is a copy of $G_{n',k'}$ thus its clique number is at most $k-2$. Therefore the second intersection has size at most $k-2$. We deduce $|K| \leq (k-2) + 2 = k$.

Let us show that $G_{n,k}$ has a Hamiltonian path. For every i , let R_i denote a Hamiltonian path of H_i , which exists by induction hypothesis. By construction, every vertex of R_i (in particular its endpoints) is adjacent to both v_{i-1} and v_i in $G_{n,k}$. Therefore $v_1 R_1 v_1 R_2 \dots v_{q-1} R_{q-1} v_q$ is a (Hamiltonian) path in G .

We proved that $G_{n,k}$ is an interval graph on at least n vertices, with clique number at most k , and with a Hamiltonian path. In order to conclude the proof, it remains to prove that G does not have an induced path longer than $n^{2/k} + 1$. Let P be an induced path of G of maximum length.

We first consider the case where $|V(P) \cap V(H_i)| \geq 2$ for some i . As for every $w \in V(H_i)$, $N(w) \setminus V(H_i) = \{v_{i-1}, v_i\}$, we deduce that $V(P) \subseteq V(H_i)$, otherwise P would not be induced. By induction hypothesis we get $|P| \leq n'^{\frac{2}{k-2}} + 1$. Observe that

$$\begin{aligned} n' &\leq \frac{n-q}{q-1} + 1 && \text{(from the definition)} \\ &\leq \frac{n}{q-1} \\ &\leq n^{\frac{k-2}{k}}. \end{aligned}$$

So $|P| \leq n^{\frac{2}{k}} + 1$, as required. We now consider the remaining case where $|V(P) \cap V(H_i)| \leq 1$ for all i . Notice that an internal vertex w of P cannot belong to H_i for some i . Indeed the only neighbors of w outside H_i are v_{i-1} and v_i , which are adjacent. We then get the three following cases:

- either $V(P)$ does not intersect any H_i , in which case $P = Q$ so $|P| \leq n^{\frac{2}{k}} + 1$;
- or there is an i such that P starts at some vertex $w \in V(H_i)$ and does not intersect H_j for all $j \neq i$, in which case $P = wv_i v_{i+1} \dots v_q$ or $P = wv_{i-1} v_{i-2} \dots v_1$ and $|P| \leq |Q| \leq n^{\frac{2}{k}} + 1$;
- or there are two integers i, j with $i < j$ such that P starts from some vertex $w_i \in V(H_i)$ and ends at some vertex w_j of $V(H_j)$, in which case $P = w_i v_i v_{i+1} \dots v_{j-1} w_j$ and again $|P| \leq |Q| \leq n^{\frac{2}{k}} + 1$. $\square$

4 Induced paths in graphs of bounded treewidth

In this section we show that the maximum function f_k such that property $(\star)$ holds for graphs of treewidth less than k is such that $f_k(n) \geq \frac{1}{4}n^{1/k}$ (Theorem 7).

Let G be a graph and let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of G . The *weight of a path* P of T is the number of vertices v of G such that T_v intersects P . The *weight of a node* x of T is the maximum weight of a path from x to a leaf minus $|\beta_{\mathcal{T}}(x)|$, and is noted $w_{\mathcal{T}}(x)$ (or $w(x)$ when there is no ambiguity on the tree representation).

It is well-known that the order of a tree is upper-bounded by a function of its height and maximum degree. The following lemma extends this statement to graphs of bounded treewidth that have a Hamiltonian path.

Lemma 22. *Let $k, w \in \mathbb{N}$ and let G be a graph that has a Hamiltonian path. If there is a tree representation of G of width less than k that has a node of weight at most w , then $|G| \leq (k+1)^{w+1} - 1$.*

Proof. Let $k \in \mathbb{N}$. The proof is by induction on w .

If $w = 0$ then for every graph G and tree representation $\mathcal{T}$ as in the statement of the lemma and x node of weight zero we have $V(G) \subseteq \beta_{\mathcal{T}}(x)$ and there are at most k vertices in $\beta_{\mathcal{T}}(x)$ so the claimed bound holds. So we now suppose that $w \geq 1$ and that the statement is true for every weight $w' < w$.

We consider a graph G with a Hamiltonian path P and a tree representation $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ of width less than k that has a node x such that $w_{\mathcal{T}}(x) \leq w$. Let us consider the graph $G \setminus \beta_{\mathcal{T}}(x)$. Removing the at most k vertices of $\beta_{\mathcal{T}}(x)$ cuts P into $t \leq k+1$ subpaths $P_1, \dots, P_t$. For each i , let G_i be the graph induced by P_i and let T_i be the union of the T_v 's for $v \in V(G_i)$. Then T_i is a subgraph of T and by Remark 17 it is connected. Observe that $\mathcal{T}_i = (T_i, \{T_v\}_{v \in V(G_i)})$ is a tree representation of G_i .

Let x_i be the node of T_i that is the closest to x in T . Let w_i be the weight of x_i in T_i and Q_i a path in T_i of maximum weight from x_i to a leaf l . Let us consider the path Q in T from x to l ; note that Q_i is a subpath of Q . So x_i belongs to Q , and by construction of T_i , $\beta_{\mathcal{T}_i}(x_i)$ is not empty and belongs to $G \setminus \beta_{\mathcal{T}}(x)$ thus there is at least one vertex in $\beta_{\mathcal{T}}(x_i) \setminus \beta_{\mathcal{T}}(x)$. This implies that $w_{\mathcal{T}}(x) \geq w_i + 1$, and thus $w_i \leq w_{\mathcal{T}}(x) - 1 \leq w - 1$.

By construction, G_i admits a Hamiltonian path and we just proved that in the tree representation $(T_i, \{T_v\}_{v \in V(G_i)})$ of width less than k there is a node x_i of weight at most $w - 1$. By induction, G_i has at most $(k+1)^{w-1} - 1$ vertices. As $V(G) = \beta_{\mathcal{T}}(x) \cup \bigcup_{i=1}^t V(G_i)$ with $t \leq k+1$ and $|\beta_{\mathcal{T}}(x)| \leq k$ we get $|G| \leq (k+1)((k+1)^{w-1} - 1) + k = (k+1)^{w+1} - 1$. $\square$

Corollary 23. *In every tree representation of width less than k of a graph of order n there is a node of weight at least $\log_{k+1}(n+1) - 1$.*

In a graph G , the *contraction* of an edge uv is the operation that creates a new vertex w adjacent to the neighbors of u and v and then deletes u and v . We say that a graph H is a *contraction* of a graph G if H can be obtained from G after a (possibly empty) sequence of edge contractions. In the sequel we use Corollary 23 to extract a graph with a long path and bounded pathwidth from a graph with a large path and bounded treewidth.

The obtained graph will be a contraction of the original one, which is interesting for us because of the following property.

Remark 24. Let H be an induced subgraph or a contraction of a graph G . If H has an induced path of order n , then so does G .

Lemma 25. *Let $k, w \in \mathbb{N}$ and let G be a graph that has a Hamiltonian path and order n . If G admits a tree representation of width less than k that has a path of weight w , then there is a contraction of G that has a Hamiltonian path and is of order w and pathwidth less than k .*

Proof. Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be the tree representation as in the statement of the lemma and R the path of T of weight w . Let P be a Hamiltonian path of G . We prove the statement by induction on the number p of vertices v of G such that $V(T_v) \cap V(R) = \emptyset$. In the case $p = 0$, for every vertex $v \in V(G)$ the subtree T_v intersects R . By the properties of tree representations, we have:

- for every $v \in V(G)$, $V(T_v) \cap V(R)$ induces a (connected) subpath of R , that we call R_v ; and
- for every $u, v \in V(G)$, T_u and T_v share a vertex if and only if they share a vertex of R .

Therefore $(R, \{R_v\}_{v \in V(G)})$ is a path representation of G . Clearly it has w vertices and width less than k hence we are done.

So we may assume in the sequel that $p > 0$ and that the statement holds for all values smaller than p . As $p > 0$ and P is a Hamiltonian path, there is an edge $uv \in E(P)$ such that T_u intersects R while T_v does not. Let us consider the graph G' obtained by contracting uv into a new vertex y and setting $T_y = T_u \cup T_v$ (which is connected as T_u and T_v intersect). We call $\mathcal{T}' = (T, \{T_z\}_{z \in V(G')})$ the corresponding tree representation. Observe that for every $t \in V(T)$, we have $\beta_{\mathcal{T}'}(t) = \beta_{\mathcal{T}}(t)$ if $t \notin T_u \cup T_v$ and $\beta_{\mathcal{T}'}(t) = \beta_{\mathcal{T}}(t) \setminus \{u, v\} \cup \{y\}$ otherwise. Therefore the width of $\mathcal{T}'$ is less than k . Also, observe that the weight of R in $\mathcal{T}'$ is still w . Applying the induction hypothesis on G' yields the desired result. $\square$

We are now ready to prove Theorem 7, that we restate here for convenience.

Theorem 7. *For every $k \in \mathbb{N}$, if G is a graph of treewidth less than k that has a path of order n , then G has an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$.*

Proof. The statement that we actually prove is the following, which implies the desired statement according to Remark 16.

For every $k, n \in \mathbb{N}$, if G is a graph of order at least n and treewidth less than k that has a Hamiltonian path, then G has an induced path of order at least $\frac{1}{4}(\log n)^{1/k}$.

Let k , n , and G be as in the statement above. The cases $n \leq 2$ or $k \leq 2$ are trivial and the case $n = k$ is handled by Theorem 6 (as then G has pathwidth at most k), so we suppose $n > k \geq 3$.

By combining Corollary 23 and Lemma 25 we obtain a contraction G' of G of order at least $\log_{k+1}(n+1)$ with pathwidth less than k and that has a Hamiltonian path.

By Theorem 6, G' admits an induced path Q of order

$$|Q| \geq \frac{1}{3}(\log_{k+1}(n+1))^{\frac{1}{k}} \geq \frac{1}{3 \cdot c}(\log(n+1))^{1/k}$$

where $c < 1.3$ is the maximum of the function $k \mapsto \log(k+1)^{1/k}$. By Remark 24 we deduce that G has an induced path of the same order and we are done. $\square$

5 Induced paths in topological minor-closed classes

In this section we use the decomposition theorem of Grohe and Marx for graphs excluding a topological minor (Theorem 38 in this paper) in order to prove Theorem 8. According to Grohe and Marx' result, such graphs admit tree decompositions where the bags⁹ are required to come from some prescribed graph classes.

We first show that $(\star)$ holds with a polylogarithmic bound for graphs from these classes (Sections 5.1 and 5.2) and then that it does too in tree representations where, intuitively, the interaction between bags is low (Section 5.3). These results are combined in Section 5.4 to finally prove Theorem 8.

5.1 Almost bounded degree graphs

Let $\Delta, k \in \mathbb{N}$. We say that a graph G has (k, Δ) -almost bounded degree if G has a set of at most k vertices whose removal yields a graph of maximum degree at most Δ . In this section we show that such graphs satisfy property $(\star)$ with a logarithmic bound (that depends on k and Δ), that is the following lemma.

Lemma 26. *For every $\Delta, k \in \mathbb{N}$ there is a constant $c \in \mathbb{R}^+$ such that if a graph with (k, Δ) -almost bounded degree has a path of order n , then it has an induced path of order at least $c \log n$.*

This result is a direct consequence of Corollary 28 (for deleting a constant number of vertices) and Corollary 30 (for graphs of bounded degree) that we prove below. They actually follow from more general statements, that we leave here as they may be useful in order to prove Conjecture 1.

Lemma 27. *Let $\mathcal{G}$ be a hereditary class of graphs such that for some $c, d \in \mathbb{R}^+$, if a graph $G \in \mathcal{G}$ has a Hamiltonian path and order n , then G has an induced path of order at least $c(\log n)^d$.*

⁹Actually the torsos, to be defined in one of the following sections.

Let $\varepsilon \in (0, 1)$ and let $\mathcal{G}_\varepsilon$ denote the class of graphs such that for every $G \in \mathcal{G}_\varepsilon$ there is a subset $X \subseteq V(G)$ of at most n^ε vertices such that $G - X \in \mathcal{G}$. Then there is a constant $c' \in \mathbb{R}^+$ depending on c , d , and ε such that if a graph $G \in \mathcal{G}_\varepsilon$ has a Hamiltonian path and order n , then G has an induced path of order at least $c'(\log n)^d$.

Proof. Let $G \in \mathcal{G}_\varepsilon$ and let P be a Hamiltonian path of G . We may assume $n \geq 3$ otherwise P is already an induced path of the desired length. Let $X \subseteq V(G)$ be such that $G - X \in \mathcal{G}$ and $1 \leq |X| \leq n^\varepsilon$, which exists by definition of $\mathcal{G}_\varepsilon$. By Remark 15, $G - X$ has a component H with a path of order at least $n' = \frac{n-|X|}{|X|+1} \geq \frac{n}{2|X|} - 1$. As $\mathcal{G}$ is hereditary, $H \in \mathcal{G}$ so it has an induced path of order at least

$$\begin{aligned} c(\log n')^d &\geq c \left(\log \left(\frac{n^{1-\varepsilon}}{2} - 1 \right) \right)^d \\ &\geq c'(\log n)^d \end{aligned}$$

for a suitable choice of the constant $c' > 0$ (depending on c , d , and ε). This path is an induced subgraph of G (by Remark 24) so we are done. $\square$

Corollary 28. Let $\mathcal{G}$ be a hereditary class of graphs such that for some $c, d \in \mathbb{R}^+$, if a graph $G \in \mathcal{G}$ has a Hamiltonian path and order n , then G has an induced path of order at least $c(\log n)^d$.

Let $k \in \mathbb{N}$ and let $\mathcal{G}_k$ denote the class of graphs such that for every $G \in \mathcal{G}_k$ there is a subset $X \subseteq V(G)$ of order at most k vertices such that $G - X \in \mathcal{G}$. Then there is a constant $c' \in \mathbb{R}^+$ depending on c , d , and k such that if a graph $G \in \mathcal{G}_k$ has a Hamiltonian path and order n , then G has an induced path of order at least $c'(\log n)^d$.

Lemma 29. Let $c \in \mathbb{R}^+$, $d \in [0, 1)$. Let $\mathcal{G}$ be the class containing every graph G with maximum degree at most $2^{c(\log |G|)^d}$. If a graph $G \in \mathcal{G}$ has a Hamiltonian path and order n , then G has an induced path of order at least $\frac{1}{c}(\log n)^{1-d}$.

Proof. Let P be a Hamiltonian path of G . Let Q be an induced path of G of maximum order and let u be one of its endpoints. Let $q = |Q|$. For every $i \in \{0, \dots, q-1\}$, let D_i denote the set of vertices at distance exactly i from u . As $D_i \subseteq N(D_{i-1})$ for every $i \in \{1, \dots, q-1\}$ we get $|D_i| \leq |D_{i-1}| \cdot 2^{c(\log n)^d}$. Therefore

$$\begin{aligned} n &= \sum_{i=0}^{q-1} |D_i| && \text{(by maximality of } Q) \\ &\leq \sum_{i=0}^{q-1} \left(2^{c(\log n)^d} \right)^i \\ &\leq \left(2^{c(\log n)^d} \right)^q \\ \text{so } q &\geq \frac{1}{c}(\log n)^{1-d}. && \square \end{aligned}$$

The following corollary about graph classes of degree bounded by a constant Δ can be obtained from the previous lemma by taking $c = \log(\Delta)$ and $d = 0$.

Corollary 30. *Let $\Delta \in \mathbb{N}$. Every graph with maximum degree at most Δ that has a Hamiltonian path and order n has an induced path of order at least $\frac{\log n}{\log \Delta}$.*

5.2 Escaping the vortices

Similar to the concept of tree representation, we can define a *cycle representation* of a graph G as a pair $\mathcal{C} = (C, \{C_v\}_{v \in V(G)})$ such that:

1. C is a cycle;
2. for every $v \in V(G)$, C_v is a connected subgraph of C ; and
3. for every edge uv of G the subgraphs C_u and C_v intersect.

The notions of *bag* and *width* of a cycle representation are defined similarly as for tree representations.

Let G_0 be a graph embedded in a surface Σ . Let C be a facial cycle of G_0 . A C -*vortex* is a cycle representation $(C, \{C_v\}_{v \in V(H)})$ of a graph H such that $V(H) \cap V(G_0) = V(C)$ and $v \in C_v$ for every $v \in V(C)$. Note that C is both a subgraph of H and the graph where the representation of H is defined.

For $g, p, a, k \in \mathbb{N}$, a graph G is (g, p, a, k) -almost-embeddable if for some set $A \subseteq V(G)$ with $|A| \leq a$ there are graphs $G_0, \dots, G_s$ with $s \leq p$ such that

1. $G - A = G_0 \cup G_1 \cup \dots \cup G_s$;
2. $G_1, \dots, G_s$ are vertex-disjoint;
3. G_0 can be embedded in a surface of Euler genus at most g ;
4. there are s pairwise vertex-disjoint facial cycles $F_1, \dots, F_s$ of G_0 in this embedding, and
5. for every $i \in \{1, \dots, s\}$, G_i has an F_i -vortex of width less than k .

This notion was introduced for the purpose of the proof of the Graph Minor Structure Theorem of Robertson and Seymour [RS03] and is also used in the decomposition theorem of Grohe and Marx on which we rely.

For $(g, 0, 0, 0)$ -almost embeddable graphs, which by definition are graphs of Euler genus at most g , property $(\star)$ is known to hold with a polylogarithmic bound as proved by Esperet et al. (Theorem 2). We use this result as a base case to show the following more general statement.

Lemma 31. *For every $g, p, a, k \in \mathbb{N}$ with $k \geq 2$ there is a constant c such that the following holds. If G is a (g, p, a, k) -almost-embeddable graph that has a Hamiltonian path and order n , then G has an induced path of order at least $c(\log n)^{\frac{1}{k}}$.*

Proof. We may assume that $n \geq 3$ otherwise the statement is trivial. Let us first assume that $a = 0$. In the case where $p = 0$ then G has Euler genus at most g and the result follows from Theorem 2. So we assume that $p \geq 1$. Let $G_0, \dots, G_s$ and $F_1, \dots, F_s$ be defined as above.

Suppose first that for some $i \in \{1, \dots, s\}$, $|F_i| \geq (k+1)(\log n + 1)$. Let $\mathcal{C} = (F_i, \{C_v\}_{v \in V(G_i)})$ be an F_i -vortex of G_i of width less than k (given by the definition of (g, p, a, k) -almost-embeddable graphs). Let $u \in V(F_i)$. The bag $\beta_{\mathcal{C}}(u)$ has size at most k so some connected component of $F_i \setminus \beta_{\mathcal{C}}(u)$ contains a subpath F of F_i of order at least $\log n$. Let $H = G[V(F)]$. This graph has F as a Hamiltonian path. Observe that by definition of F , no model (in $\mathcal{C}$) of a vertex of F contains u . In other words, the models of the vertices of F are all subpaths of $F_i \setminus \{u\}$. Hence $(F_i \setminus \{u\}, \{C_v\}_{v \in V(H)})$ is a path representation of width less than k of H . By Theorem 6 we get an induced path of order at least $\frac{1}{3}(\log n)^{1/k}$ in H hence in G . We are not completely done yet but let us now consider the second case before concluding.

In the second case we assume that $|F_i| < (k+1)(\log n + 1)$ for every $i \in \{1, \dots, s\}$. This implies that $|G_i| < k(k+1)(\log n + 1)$. Let $X = \bigcup_{i=1}^s V(G_i)$, then $1 \leq |X| < pk(k+1)(\log n + 1)$.

By Remark 15, $G - X$ has a connected component G' that has a Hamiltonian path and order n' where $n' \geq \frac{n}{2|X|} - 1$. So $n' \geq n^\varepsilon$ for some constant $\varepsilon \in (0, 1)$ that depends on p and k only.

Because it is a subgraph of G_0 , the graph G' has Euler genus at most g . Applying Theorem 2 there is a constant c_g depending on g only such that G' (hence G) has an induced path of order q at least

$$\begin{aligned} q &\geq c_g \sqrt{\log n'} \\ &\geq c_g \sqrt{\varepsilon} \sqrt{\log n} \\ &\geq c_g \sqrt{\varepsilon} (\log n)^{1/k} \end{aligned} \quad \text{as } k \geq 2.$$

Let $c = \min(1/3, c_g \sqrt{\varepsilon})$. In both cases we obtained an induced path of order at least $c(\log n)^{1/k}$, as claimed.

The case where $a \geq 1$ follows from Corollary 28 applied to the case where $a = 0$. This concludes the proof. $\square$

5.3 Representations of bounded adhesion

In this section we build upon the ideas developed in Sections 3 and 4 to show that, roughly, if $\mathcal{G}$ is a class of graphs where $(\star)$ holds with a polylogarithmic bound then the same can be said of graphs obtained by gluing together in a tree-like fashion graphs from $\mathcal{G}$ (Lemma 36). This is a crucial step towards the proof of Theorem 8.

Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of a graph G . The *adhesion set* of an edge tt' of T is defined as the following subset of $V(G)$:

$$\text{adh}_{\mathcal{T}}(tt') = \{v \in V(G) \mid \{t, t'\} \subseteq V(T_v)\}.$$

Equivalently, it can be defined as the intersection of the bags at t and t' , that is $\text{adh}_{\mathcal{T}}(tt') = \beta_{\mathcal{T}}(t) \cap \beta_{\mathcal{T}}(t')$. We drop the subscript when it is clear from the context. The *adhesion* of $\mathcal{T}$ is the maximum size of the adhesion set of an edge of T . For every $t \in V(T)$, the *torso* of $\mathcal{T}$ at t is the graph obtained from $G[\beta(t)]$ by adding all edges uv such that $u, v \in \text{adh}_{\mathcal{T}}(tt')$ for some neighbor t' of t . If for every $t \in V(T)$, the torso of $\mathcal{T}$ at t belongs to some graph class $\mathcal{G}$, we say that $\mathcal{T}$ has torsos from $\mathcal{G}$.

Torsos and adhesion sets provide two different ways to restrict tree representations. Observe that graphs of treewidth (respectively pathwidth) less than k are simply graphs that admit a tree representation (respectively path representation) with torsos from the class of graphs of order at most k . For every graph class $\mathcal{G}$ and integer a , we respectively denote by $\text{PR}(\mathcal{G}, a)$ and $\text{TR}(\mathcal{G}, a)$ the class of graphs that admit a path representation or a tree representation with torsos from $\mathcal{G}$ and adhesion less than a . We denote by $\text{TR}(\mathcal{G})$ the class of graphs that admit a tree representation with torsos from $\mathcal{G}$ and no restriction on the adhesion. Notice that if the graphs in $\mathcal{G}$ have cliques of bounded order, the adhesion of tree representations of graphs from $\text{TR}(\mathcal{G})$ is implicitly bounded (i.e. $\text{TR}(\mathcal{G}) \subseteq \text{TR}(\mathcal{G}, a)$ for some $a \in \mathbb{N}$).

The next remark easily follows from the definition of tree representations.

Remark 32. Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of a graph G , let $tt' \in E(T)$ and let F, F' be the two connected components of $T \setminus \{tt'\}$. Let u and u' be such that $V(T_u) \cap V(F) \neq \emptyset$ and $V(T_{u'}) \cap V(F') \neq \emptyset$. Then either one of u, u' belongs to $\text{adh}_{\mathcal{T}}(tt')$, or $G \setminus \text{adh}_{\mathcal{T}}(tt')$ has no path from u to u' .

For the purpose of the proof of Lemma 36 we need to relate the order of a graph with the length of a path where it is represented. It is not true in general that the existence of a path representation with a long path implies that the represented graph is large; for instance all the bags in this representation could be identical and small. We show below (Lemma 33) that such a statement holds if we require the considered path representation to satisfy an extra property, being varied, that we define now.

Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a tree representation of a graph G . For an edge tt' in T , the tree representation $(T', \{T'_v\}_{v \in V(G)})$ obtained from $\mathcal{T}$ by *contracting* tt' is defined as follows:

- T' is obtained from T by contracting tt' ;
- for every $v \in V(G)$, $T'_v = T_v$ if $\{t, t'\} \cap T_v = \emptyset$ and otherwise $T'_v = T_v \setminus \{t, t'\} \cup \{t''\}$.

Intuitively, we merge the nodes t and t' both in the tree of the representation and in the models of the vertices.

Let us say that $\mathcal{T}$ is *varied* if no bag is a subset of a neighboring bag, i.e. for every $tt' \in E(T)$, $\beta(t) \not\subseteq \beta(t')$. In particular, unless G has no vertex, no bag is empty.

Given a tree representation $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$, it is possible to produce a varied tree representation by iteratively contracting in $\mathcal{T}$ the edges tt' such that $\beta(t) \subseteq \beta(t')$. Observe that this process changes neither the width or the adhesion of $\mathcal{T}$ nor the fact that it has bags from some specific class of graphs.

Lemma 33. *Let G be a graph on at least one vertex. If $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ is a varied path representation of G , then $|G| \geq |R|$.*

Proof. We prove the following statement by induction on ℓ .

For every $\ell \in \mathbb{N}^+$, if a graph G on at least one vertex admits a varied path representation on a path of order ℓ , then $|G| \geq \ell$.

The case $\ell = 1$ is trivial as we require that G is not empty. So let us assume that $\ell > 1$ and that the statement holds for smaller values. Let G be a graph that admits a varied path representation $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ with $|R| = \ell$, and let $r_1 \dots r_\ell$ be the vertices of R in the order of the path. Let $G^- = G[\beta(r_1) \cup \dots \cup \beta(r_{\ell-1})]$. As $\mathcal{R}$ is varied, $\beta(r_1) \neq \emptyset$ so G^- has at least one vertex. Observe that it admits a varied path representation with $\ell - 1$ nodes (for instance $(R \setminus \{r_\ell\}, \{R_v\}_{v \in V(G^-)})$), so by induction it has at least $\ell - 1$ vertices. As $\mathcal{R}$ is varied, there is a vertex $v \in \beta(r_\ell)$ that does not belong to $\beta(r_{\ell-1})$. This vertex does not belong to G^- either (as R_v is connected) so $|G| \geq \ell$, as claimed. $\square$

The above lemma allows us to prove the following variant of Theorem 6 on varied path representations of bounded adhesion.

Lemma 34. *Let G be a graph that has a Hamiltonian path. If G admits a varied path representation with adhesion less than a and at least ℓ nodes, then G has an induced path of order at least $\frac{1}{3}\ell^{\frac{1}{2a}}$.*

Proof. Let $\mathcal{R} = (R, \{R_v\}_{v \in V(G)})$ be a path representation as in the statement. For every $r \in V(R)$, let Z_r be a subset of $V(G)$ of minimum size such that for every neighbor r' of r , $\text{adh}(rr') \subseteq Z_r$ and there is a vertex $v \in Z_r$ that does not belong to $\beta(r')$. Let us show that this set is well-defined and small. Either the bag at r has a vertex v that does not appear in the bags of any of its neighbors, in which case $\{v\} \cup \bigcup_{r' \in N(r)} \text{adh}(rr')$ satisfies the above properties, or it does not and then $\beta(r)$ is suitable, as $\mathcal{R}$ is varied. Observe that in this case $\beta(r) = \bigcup_{r' \in N(r)} \text{adh}(rr')$. Recall that in R the vertex r has up to two neighbors. So in both cases we have $|Z_r| \leq 2(a - 1) + 1 = 2a - 1$.

Let P be a Hamiltonian path of G . An edge of P is called *superfluous* if it has at least one endpoint outside $\bigcup_{r \in V(R)} Z_r$.

Remark 35. Suppose that xy is a superfluous edge and let G' be the graph obtained by contracting xy into a new vertex z . Then $\mathcal{R}' = (R, \{R_v\}_{v \in V(G')})$ is a path representation of G' (with $R_z = R_x \cup R_y$). We show that additionally $\mathcal{R}'$ is a varied representation. Towards a contradiction, let us assume that there are $r, r' \in V(R)$ such that $\beta_{\mathcal{R}'}(r) \subseteq \beta_{\mathcal{R}'}(r')$. Then for one of x and y , say x , we have $\beta_{\mathcal{R}}(r) \setminus \beta_{\mathcal{R}}(r') = \{x\}$, which implies $y \in \text{adh}_{\mathcal{R}}(rr')$ (as xy is an edge). We just proved that $x, y \in Z_r$, which is a contradiction with the choice of xy . So $\mathcal{R}'$ is varied.

Let H be the graph obtained from G after iteratively contracting all the superfluous edges and let $\mathcal{R}_H = (R, \{R_v\}_{v \in V(H)})$ be the corresponding varied path representation constructed as in the above remark. By Lemma 33, $|H| \geq |R| \geq \ell$. All the edges that were contracted did belong to P so H has a Hamiltonian path. No edge of H is superfluous

(by definition) so for every $r \in V(R)$, $\beta_{\mathcal{R}_H}(r) = Z_r$. Therefore $\mathcal{R}_H$ is a path representation of H of width at most $2a - 1$. Applying Theorem 6 to H we get an induced path of order at least $\frac{1}{3}\ell^{\frac{1}{2a}}$. As H is a contraction of G , such a path also exists in G (Remark 24) hence we are done. $\square$

A graph class is said to be *closed under taking subgraphs* if every subgraph of a graph of the class also belongs to the class. We are now ready to prove the main result of this section.

Lemma 36. *Let $a \in \mathbb{N}^+$, $c \in (0, 1/3]$ and $d \in (0, 1]$. Let $\mathcal{G}$ be a class of graphs that is closed under taking subgraphs and such that for every $n \in \mathbb{N}$, every $G \in \mathcal{G}$ that has a Hamiltonian path and order n has an induced path of order at least $c(\log n)^d$.*

If a graph G of $\text{TR}(\mathcal{G}, a)$ has a Hamiltonian path and order n , then it has an induced path of order at least $c(\log n)^{\frac{1}{4a+\frac{1}{d}}}$.

Proof. Let $\mathcal{T} = (T, \{T_v\}_{v \in V(G)})$ be a varied tree representation of G witnessing that $G \in \text{TR}(\mathcal{G}, a)$. Let us fix $\varepsilon = \frac{4ad}{4ad+1} < 1$. Observe that we may assume

$$c(\log n)^{\frac{1}{4a+\frac{1}{d}}} > 2 \quad (1)$$

as otherwise the statement holds trivially. Let P be the Hamiltonian path of G .

Let us first suppose that T has a node t such that the bag $\beta(t)$ has order at least $n^{\frac{1}{(\log n)^\varepsilon}}$. Let H denote the graph obtained from G by iteratively contracting every edge of P that has at least one endpoint outside $\beta(t)$. By construction, H is a supergraph of $G[\beta(t)]$ of the same order and every edge $uv \in E(H) \setminus E(G[\beta(t)])$ is such that $u, v \in \text{adh}_{\mathcal{T}}(tt')$ for some neighbor t' of t . Therefore, H is a subgraph of the torso of $\mathcal{T}$ at t . As $\mathcal{G}$ is subgraph-closed, we deduce $H \in \mathcal{G}$. Besides, these contractions were applied to edges of P so they yield a Hamiltonian path in H . By the properties of $\mathcal{G}$ we deduce that H has an induced path of order at least

$$\begin{aligned} c \left(\log n^{\frac{1}{(\log n)^\varepsilon}} \right)^d &= c(\log n)^{(1-\varepsilon)d} \\ &= c(\log n)^{\frac{1}{4a+\frac{1}{d}}}. \end{aligned}$$

Such a path also exists in G (Remark 24) so we are done.

So we may assume now that at every node t of T we have

$$|\beta(t)| < n^{\frac{1}{(\log n)^\varepsilon}}, \quad (2)$$

which implies

$$|T| \geq \frac{n}{n^{\frac{1}{(\log n)^\varepsilon}}} = n^{1-\frac{1}{(\log n)^\varepsilon}}. \quad (3)$$

Let $t \in V(T)$. As $\mathcal{T}$ is varied, for every neighbor t' of t there is a vertex $v_{t'} \in \beta(t')$ of G such that $v_{t'} \notin \beta(t)$. Let $G_{t'}$ denote the connected component of $G \setminus \beta(t)$ that contains

$v_{t'}$. Then for every neighbor $t'' \neq t'$ of t , the components $G_{t'}$ and $G_{t''}$ are distinct, by the properties of tree representations. So $G \setminus \beta(t)$ has at least $\deg_T(t)$ connected components. By (2) and Remark 15, $G \setminus \beta(t)$ has at most $|\beta(t)| + 1$ connected components. We deduce that the maximum degree Δ of T is bounded as follows

$$\begin{aligned}\Delta &\leq n^{\frac{1}{(\log n)^\varepsilon}} + 1 \\ &\leq 2n^{\frac{1}{(\log n)^\varepsilon}}.\end{aligned}\tag{4}$$

Let R denote a path of maximum order in T . From the classic inequality $|T| \leq \Delta^{|R|-1}$ we get

$$\begin{aligned}|R| - 1 &\geq \frac{\log |T|}{\log \Delta} \\ &\geq \frac{\log n - (\log n)^{1-\varepsilon}}{(\log n)^{1-\varepsilon} + 1} && \text{from (3) and (4)} \\ &\geq \frac{1}{2}((\log n)^\varepsilon - 1) && \text{from (1)} \\ \text{so } |R| &\geq \frac{1}{2}(\log n)^\varepsilon \\ &\geq (\log n)^{\varepsilon/2} && \text{from (1)}.\end{aligned}$$

As in the first part of the proof we iteratively contract the edges of P that do not have both endpoints in $\bigcup_{t \in V(R)} \beta(t)$ in order to produce a graph H that has a Hamiltonian path and a varied path representation $(R, \{R_v\}_{v \in V(H)})$ with adhesion less than a . We can now apply Lemma 34 to get an induced path of order at least

$$\begin{aligned}\frac{1}{3}|R|^{\frac{1}{2a}} &\geq \frac{1}{3}(\log n)^{\frac{\varepsilon}{4a}} \\ &\geq c(\log n)^{\frac{1}{4a+\frac{1}{d}}} && \text{as } c \leq \frac{1}{3}.\end{aligned}$$

Again such a path also exists in G and we are done. □

5.4 Piecing things together

We are now ready to prove the following theorem from which will follow Theorem 8.

Theorem 37. *Let $k \in \mathbb{N}$ and let $\mathcal{G}_k$ denote the class of graphs that either are (k, k, k, k) -almost embeddable or have (k, k) -almost bounded degree. There are constants $c \in \mathbb{R}^+$ and $d \in (0, 1)$ such that if a graph $G \in \text{TR}(\mathcal{G}_k)$ has a Hamiltonian path and order n , then G has an induced path of order at least $c(\log n)^d$.*

Proof. Observe that a (k, k) -almost bounded degree graph does not contain a clique of order $2k + 2$. Also there is a $k' \in \mathbb{N}$ such that no (k, k, k, k) -almost embeddable graph contains a clique of order k' (see for instance [DMW17, Lemma 21] for a linear upper-bound in terms of k). So for $a = \max(2k + 2, k')$ we have $\text{TR}(\mathcal{G}_k) \subseteq \text{TR}(\mathcal{G}_k, a)$. By

Lemma 31 and Lemma 26 the class $\mathcal{G}_k$ satisfies $(\star)$ with the function $n \mapsto c'(\log n)^{d'}$ for some constants $c' \in \mathbb{R}^+$ and $d' \in (0, 1)$ depending on k . Also, notice that $\mathcal{G}_k$ is closed under taking subgraphs. Together with Lemma 36 this implies that $\text{TR}(\mathcal{G}_k, a)$ (hence $\text{TR}(\mathcal{G}_k)$) satisfies $(\star)$ with the function $n \mapsto c(\log n)^d$ for some constants $c \in \mathbb{R}^+$ and $d \in (0, 1)$ depending on k . $\square$

Theorem 8 is a consequence of Theorem 37 and the structure theorem of Grohe and Marx for graphs excluding a topological minor, stated hereafter in the setting of tree representations.

Theorem 38 ([GM15]). *For every graph H there is an integer $k \in \mathbb{N}$ such that every graph not containing H as a topological minor has a tree representation $(T, \{T_v\}_{v \in V(G)})$ such that for every $t \in V(T)$ the torso at t is either (k, k, k, k) -almost embeddable or has (k, k) -almost bounded degree.*

6 Open problems

A first direction for future work is to investigate how widely the results proved in this paper could be generalized. What are the most general graph classes where the bound for property $(\star)$ is (at least) polylogarithmic? We recall that it was conjectured in [ELM17] that it is the case for k -degenerate graphs, for every $k \in \mathbb{N}$ (Conjecture 1). Note that the bounds we obtained in Theorem 6 and Corollary 12 are polynomial. An interesting task could be to characterize hereditary classes where such a property holds. Also, as mentioned in the introduction, Esperet et al. proved that in k -trees (i.e. edge-maximal graphs of treewidth k) the property $(\star)$ holds with a $\Omega(\log n)$ bound, while their upper-bound for graphs of treewidth at most k (recalled in the table below) shows that such a bound where k does not appear in the exponent of $\log n$ does not hold for graphs of treewidth at most k . This suggests that our results could be improved in the restricted setting of edge-maximal graphs from the considered classes.

Finally, a natural research direction about this problem is to obtain tight bounds for our theorems, especially in the cases of bounded pathwidth or treewidth, for which we recall below the known bounds.

class	lower-bound	upper-bound
graphs of pw $< k$	$\frac{1}{3}n^{1/k}$ (Th. 6)	$n^{2/k} + 1$ (Th. 21)
graphs of tw $< k$	$\frac{1}{4}(\log n)^{1/k}$ (Th. 7)	$(k+1)(\log n)^{2/(k-1)}$ [ELM17]

To the best of our knowledge, this question is also open for planar graphs (and more generally graphs of bounded Euler genus) with the current best lower- and upper-bounds on the function f of property $(\star)$, both due to [ELM17], being

$$\left(\frac{1}{2\sqrt{6}} - o(1) \right) \sqrt{\log n} \leq f(n) \leq \frac{3 \log n}{\log \log n}.$$

References

- [ALR19] Aistis Atminas, Vadim Lozin, and Igor Razgon. Graphs without large bicliques and well-quasi-orderability by the induced subgraph relation. *Journal of Combinatorics*, 10(2):327–337, 2019.
- [AV00] Jorge Arocha and Pilar Valencia. Long induced paths in 3-connected planar graphs. *Discussiones Mathematicae Graph Theory*, 20(1):105–107, 2000.
- [CO00] Bruno Courcelle and Stephan Olariu. Upper bounds to the clique width of graphs. *Discrete Applied Mathematics*, 101(1):77–114, 2000.
- [CY02] Guantao Chen and Xingxing Yu. Long cycles in 3-connected graphs. *Journal of Combinatorial Theory, Series B*, 86(1):80–99, 2002.
- [DLM16] Emilio Di Giacomo, Giuseppe Liotta, and Tamara Mchedlidze. Lower and upper bounds for long induced paths in 3-connected planar graphs. *Theoretical Computer Science*, 636:47–55, 2016.
- [DMW17] Vida Dujmović, Pat Morin, and David R. Wood. Layered separators in minor-closed graph classes with applications. *Journal of Combinatorial Theory, Series B*, 127:111–147, 2017.
- [ELM17] Louis Esperet, Laetitia Lemoine, and Frédéric Maffray. Long induced paths in graphs. *European Journal of Combinatorics*, 62:1–14, 2017.
- [Gav74] Fănică Gavril. The intersection graphs of subtrees in trees are exactly the chordal graphs. *Journal of Combinatorial Theory, Series B*, 16(1):47–56, 1974.
- [GGL95] Ron Graham, Martin Grötschel, and László Lovász. *Handbook of Combinatorics*, volume 1. Elsevier, 1995.
- [GLP⁺21] Peter Gartland, Daniel Lokshtanov, Marcin Pilipczuk, Michał Pilipczuk, and Paweł Rzażewski. Finding large induced sparse subgraphs in $C_{>t}$ -free graphs in quasipolynomial time. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 330–341, 2021.
- [GM15] Martin Grohe and Dániel Marx. Structure theorem and isomorphism test for graphs with excluded topological subgraphs. *SIAM Journal on Computing*, 44(1):114–159, 2015.
- [GRS82] F Galvin, I Rival, and B Sands. A Ramsey-type theorem for traceable graphs. *Journal of Combinatorial Theory, Series B*, 33(1):7–16, 1982.
- [GW00] Frank Gurski and Egon Wanke. The tree-width of clique-width bounded graphs without $K_{n,n}$. In Ulrik Brandes and Dorothea Wagner, editors, *Graph-Theoretic Concepts in Computer Science*, pages 196–205, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.
- [NDM12] Jaroslav Nešetřil and Patrice Ossona De Mendez. *Sparsity: graphs, structures, and algorithms*, volume 28. Springer Science & Business Media, 2012.

- [PT99] Andrzej Proskurowski and Jan Arne Telle. Classes of graphs with restricted interval models. *Discrete Mathematics & Theoretical Computer Science*, Vol. 3 no. 4, January 1999.
- [RS03] Neil Robertson and Paul D Seymour. Graph minors. XVI. excluding a non-planar graph. *Journal of Combinatorial Theory, Series B*, 89(1):43–76, 2003.