Minimal hereditary graph classes of unbounded clique-width

by

Daniel Cocks

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the
School of Mathematics and Statistics
STEM Faculty
The Open University

August 2024

Declaration of Authorship

This thesis is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Signed: Daniel Cocks

Date: 5th August 2024

Inspiration thanks to Alfred, Lord Tennyson, taken from 'Ulysses', 1833

Death closes all: but something ere the end,
Some work of noble note, may yet be done,
Not unbecoming men that strove with Gods.
The lights begin to twinkle from the rocks:
The long day wanes: the slow moon climbs: the deep
Moans round with many voices. Come, my friends,
'T is not too late to seek a newer world.
Push off, and sitting well in order smite
The sounding furrows; for my purpose holds
To sail beyond the sunset, and the baths
Of all the western stars, until I die.
It may be that the gulfs will wash us down:
It may be we shall touch the Happy Isles,
And see the great Achilles, whom we knew.
Tho' much is taken, much abides; and tho'
We are not now that strength which in old days
Moved earth and heaven, that which we are, we are;
One equal temper of heroic hearts,
Made weak by time and fate, but strong in will
To strive, to seek, to find, and not to yield.

Abstract

In the study of graphs, clique-width is a parameter that has received much attention due its significance in the tractability of algorithms on certain classes of graph. Of particular interest are hereditary graph classes, those classes closed under taking induced subgraphs. A number of *minimal hereditary graph classes of unbounded clique-width* (abbreviated to *minimal classes*) have recently been identified; that is, classes containing graphs with arbitrarily large clique-width but where every proper hereditary subclass has bounded clique-width. There are also hereditary classes of unbounded clique-width that do not contain a minimal subclass, but instead contain graph structures known as t-*basic obstructions* to bounded clique-width for arbitrarily large t. These graphs form a sequence known as an *antichain of unbounded clique-width*. We identify many new minimal classes and place all known minimal classes inside two 'frameworks'. We also identify new t-basic obstructions to bounded clique-width.

In Chapters 2 and 3 we create our first framework for dense minimal classes, consisting of graph classes constructed by taking the finite induced subgraphs of an infinite graph $\mathcal{P}^{\delta}$ whose vertices form a two-dimensional array and whose edges are defined by three objects, denoted as a triple $\delta = (\alpha, \beta, \gamma)$. We introduce new methods to the study of clique-width, and identify uncountably many new minimal classes in the framework.

In sparse classes clique-width is unbounded if and only if the (widely studied) parameter, tree-width, is unbounded. In Chapter 4 we identify a new t-basic obstruction, a t-*sail*. We construct 'path-star' graph classes defined by a *nested* word, with a recursive structure, in which a graph has large tree-width if and only if it contains a large t-sail. We show that these classes are infinitely defined and do not contain a minimal subclass.

In Chapter 5 we create an alternative framework for minimal classes to the one developed in Chapters 2 and 3, containing 'path-clique' graph classes consisting of the finite induced subgraphs of an infinite graph created from the symmetric difference of edges between an infinite path and a partition of the path vertices forming infinite cliques or independent sets that are complete or anti-complete to each other. We identify another uncountable family of minimal classes different to those from the first framework.

In Chapter 6 we identify a new t-basic obstruction – *a* t-*clipper*. We show that a graph in the class of permutation-partition graphs has large clique-width if and only if it contains a large t-clipper. We also identify other likely t-basic obstructions.

Acknowledgements

I belong to the class of 2020 – that ill-fated year of Covid. Across the width of the land we became a giant independent set of nodes, each isolated in our own neighbourhood — contact was minimal (and then only remotely). It was more than a year before we were induced to campus to be adjacent to one another. However, when I finally met the OU Maths and Stats staff and students, no algorithm was needed to prove that friendliness is not hereditary -– they were welcoming with no sign of cliques.

Although my praise may appear sparse, this would be anticomplete without acknowledging the support of the mathematical set who gave this dense but tractable student an edge with very helpful anonymous evaluation of my modest output.

Most thanks of all go to Robert Brignall who provided an unbounded amount of time and wise advice and to my wife Jane who shared with me an uncountable number of cheese scones.

Thanks go also to the antichain of family and friends who listened politely to an explanation of my research without going too glassy-eyed.

Lastly, whichever permutation of factors you take, a key parameter is money, so to UK taxpayers who funded this research, you have my infinite gratitude.

Publications

Much of the content of this thesis has previously been published in the form of papers, as follows:

1. The material on the framework for minimal dense classes in Chapters 2 and 3 has been published in [14], co-authored with R. Brignall.

2. The material on uncountably many minimal classes and classes defined by recurrent words in Chapter 3 has been published in [13], co-authored with R. Brignall.

3. The content of Chapter 4 on sparse graph classes has been published in [16].

Contents

Chapter 1

Introduction and background

1.1 Evaluation problems defined on a graph

We say that an algorithm to solve a problem, with input size n, is intractable if it takes an exponential amount of time or worse, estimated as the number of steps required as a function of n. Many problems can be represented as networks, or graphs, and therefore it is natural to consider the mathematical properties of graphs in studying the tractability of algorithms to solve both theoretical and real-world problems. Researchers are typically interested in algorithms to solve 'evaluation' problems (sometimes referred to as 'decision' or 'membership' problems): that is, given a graph G on n vertices and some property, does the graph possess the property?

When applied to an arbitrary graph, an evaluation problem may carry no guarantee of being solved within a reasonable timeframe (if it can be solved at all). However, it is often the case that the same problem has an efficient solution when applied to graphs from some restricted collection of graphs, or graph class.

1.2 Definitions

1.2.1 Graphs

Throughout this thesis we mainly use standard graph theoretic notation from Diestel [30] with the following variations:

If vertex u is adjacent to vertex v we write $u \sim v$ and if u is not adjacent to v we write $u \not\sim v$. We denote $N(v)$ as the neighbourhood of a vertex v, that is, the set of vertices adjacent to v. We denote a clique with r vertices as K_r and an independent set of r vertices as $\overline{K_r}$. Two subsets X and Y of V are said to be *complete* if every vertex of X is adjacent to every vertex of Y. Likewise, X and Y are said to be *anticomplete* if there is no edge of G between X and Y.

A graph $G = (V, E)$ is *bipartite* if it admits a partition of V into two independent sets, X and Y, such that every edge has an end in X and an end in Y. A graph $G = (V, E)$ is *complete bipartite* or is called a *biclique* if it admits a partition of V into two independent sets X and Y that are complete. We denote this by $K_{r,s}$ where $|X| = r$ and $|Y| = s$. A graph $G = (V, E)$ is a *complete split* graph if it admits a partition of V into two sets X and Y that are complete, with X a clique and Y an independent set. A graph $G = (V, E)$ is a *half graph* if it admits a partition of V into two independent sets $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_n\}$ where $x_i \sim y_j$ whenever $i \geqslant j$.

A *star* S_k is the complete bipartite graph $K_{1,k}$: a tree with one internal vertex, which we will refer to as the *star-vertex*, and k *leaves*.

A k-*cycle* is a closed path with k vertices. A k-cycle has a *chord* if two of its k vertices are joined by an edge which is not itself part of the cycle. A *hole* is a chordless cycle of length at least 4.

The *complement* of a graph G is denoted $\overline{G}$ (that is, $\overline{G}$ has the same vertices as G such that two distinct vertices of $\overline{G}$ are adjacent if and only if they are not adjacent in G). The *bipartite complement* of a bipartite graph G with partite sets X and Y is denoted $\widetilde{G}$ (that is, $\widetilde{G}$ is a bipartite graph with the same partite sets X and Y as G, where $x \in X$ is adjacent to $y \in Y$ if and only if x is not adjacent to y in G). The *line* graph of a graph G is denoted $L(G)$ (that is, $L(G)$ is a graph with vertices corresponding to the edges of G such that two vertices are adjacent if and only if they are adjacent as edges in G).

We will use the notation $H \leqslant G$ to denote graph H is an *induced subgraph* of graph G. If graph G does not contain an induced subgraph isomorphic to H we say that G is H-*free*. A class of graphs is *hereditary* if it is closed under taking induced subgraphs. The notation $\mathcal{C} \subseteq \mathcal{G}$ denotes that $\mathcal{C}$ is a *hereditary subclass* of hereditary graph class $\mathcal{G}$ ($\mathcal{C} \subsetneq \mathcal{G}$ for a proper subclass). If no graph in $\mathcal{C}$ contains an induced subgraph H this is denoted $\mathcal{C} \subseteq \mathsf{Free}(H)$.

An *embedding* of graph H in graph G is an injective map $\phi : V(H) \to V(G)$ such that the subgraph of G induced by the vertices $\phi(V(H))$ is isomorphic to H. In other words, $vw \in E(H)$ if and only if $\phi(v)\phi(w) \in E(G)$. If H is an induced subgraph of G then this can be witnessed by one or more embeddings.

All graphs are finite, simple and loopless, unless otherwise stated.

1.2.2 Words

We make significant use of combinatorics on words for which our main reference work is Pytheas Fogg [31].

We refer to a (finite or infinite) sequence of letters chosen from a finite or infinite alphabet as a *word*. Here we only use $\mathbb{N}$ or a subset of $\mathbb{N} \cup \{0\}$ for our alphabet. We denote by ω_i the i-th letter of the word ω. A *factor* of ω is a contiguous subword $\omega_{[i,j]}$ being the sequence of letters from the i-th to the j-th letter of ω. If a is a letter from the alphabet we will denote a^∞ as the infinite word $aaa\ldots$, and if $a_1 \ldots a_n$ is a finite sequence of letters from the alphabet then we will denote $(a_1 \ldots a_n)^\infty$ as the infinite word consisting of the infinite repetition of this factor.

The *length* of a word (or factor) is the number of letters the word contains, and the *distance* between the i-th and j-th letter in a word is $|i - j|$.

An infinite word ω is *recurrent* if each of its factors occurs in it infinitely many times. We say that ω is *almost periodic* (sometimes called *uniformly recurrent* or *minimal*) if for each factor $\omega_{[i,j]}$ of ω there exists a constant $\mathcal{L}(\omega_{[i,j]})$ such that every factor of ω of length at least $\mathcal{L}(\omega_{[i,j]})$ contains $\omega_{[i,j]}$ as a factor. Finally, ω is *periodic* if there is a positive integer p such that $\omega_k = \omega_{k+p}$ for all k. Clearly, every periodic word is almost periodic, and every almost periodic word is recurrent.

Given a word α over an alphabet $\mathcal{A}$, and a sub-alphabet $\mathcal{S} \subset \mathcal{A}$, the *subword* of α restricted to $\mathcal{S}$ is the word derived from α by deleting all letters not in $\mathcal{S}$ and concatenating the remaining factors in the same order as they appear in α. We denote this subword as $\alpha^{\mathcal{S}}$.

1.3 Graph parameters

One particularly rich source of graph classes where evaluation problems can be solved efficiently are those in which some graph parameter is bounded. In general, a graph parameter is a positive-integer-valued measure of some aspect of complexity of a graph. The higher the value, the more complex the graph. There are many such parameters, but here we will only use *tree-width*, *clique-width* (and the related *linear clique-width*) and *rank-width*.

1.3.1 Tree-width

Definition 1.1. Let $G = (V, E)$ be a graph , T a tree, and let $\mathcal{V} = (V_t)_{t \in T}$ be a family of vertex sets $V_t \subseteq V(G)$ (called *bags*) indexed by the vertices t of T. The pair $(T, \mathcal{V})$ is called a *tree-decomposition* of G if it satisfies the following three conditions:

1. Every vertex of G is in at least one of the bags V_t,
2. If $(u, v) \in E$, then u and v are together in some bag,
3. for all $v \in V$, the graph induced by the bags containing v is connected in T.

The *width* of a tree-decomposition is the maximum bag size minus 1. The *tree-width* of G, denoted by $tw(G)$, is the least width of any tree-decomposition of G.

A key characteristic of tree-width is the following:

Lemma 1.2 ([30]). *If* H *is a minor of* G *then* $tw(H) \leqslant tw(G)$.

Tree-width measures how 'tree-like' a graph is, in the sense that the lower the tree-width the more like a tree is the graph (thus, graphs with tree-width 1 are exactly the trees and the forests and graphs with tree-width at most 2 are the series-parallel graphs). Tree-width is the subject of Robertson and Seymour's Grid Minor Theorem [57] which states that every graph of large enough tree-width must contain a minor isomorphic to a large grid. This result is central to Robertson and Seymour's sequence of over twenty papers, which collectively establish a deep and fundamental connection between the theory of graph minors and algorithms, culminating in the Graph Minor Theorem [58]. Reinhard Diestel wrote that this body of work 'dwarfs any other result in graph theory and may doubtless be counted among the deepest theorems that mathematics has to offer'. As a result of Lemma 1.2, tree-width has largely been associated with minor-closed graph classes (that is, the class includes all graphs formed from another in the class by a process of vertex or edge deletion, or edge 'contraction'). Such classes include forests and planar graphs.

1.3.2 Clique-width and linear clique-width

Clique-width is a graph width parameter introduced by Courcelle, Engelfriet and Rozenberg in the 1990s [19] as a generalisation of tree-width.

Definition 1.3. The clique-width of a graph is denoted $cw(G)$ and is defined as the minimum number of labels needed to construct G by means of the following four graph operations:

(a) creation of a new vertex v with label i (denoted $i(v)$),

(b) adding an edge between every vertex labelled i and every vertex labelled j for distinct i and j (denoted $\eta_{i,j}$),

(c) giving all vertices labelled i the label j (denoted $\rho_{i\to j}$), and

(d) taking the disjoint union of two previously-constructed labelled graphs G and H, one of which may be empty (denoted $G \oplus H$).

The *linear clique-width* of a graph G denoted $lcw(G)$ is the minimum number of labels required to construct G by means of four operations, being (a), (b), (c) above plus

(d′) taking the disjoint union of two previously-constructed labelled graphs G and H, one of which is a single labelled vertex v (denoted $G \oplus v$) or no vertex (denoted $G \oplus \emptyset$).

Every graph can be defined by an algebraic expression τ using the four operations above, which we will refer to as a *(linear) clique-width expression*. This expression is called a k-*expression* if it uses k different labels.

A clique-width expression τ defining G can be represented as a rooted binary tree, tree(τ), whose leaves correspond to the operations of vertex creation, the internal nodes correspond to the $\oplus$-operation, and the root is associated with G. The operations η and ρ are assigned in the appropriate sequence along the respective edges of tree(τ). The tree is binary since each $\oplus$-operation brings together at most two previously constructed graphs. Also, it can be observed that an $\oplus$-vertex represents a subgraph of G but not usually an induced subgraph since there may still be edges to be created by η operations.

In the case of a linear clique-width expression the tree becomes a *caterpillar tree*, that is, a tree that becomes a path after the removal of the leaves.

As already noted, a class of graphs is hereditary if it is closed under taking induced subgraphs (that is, the class includes all graphs formed from another in the class by a process of vertex deletion only). As well as being conceptually simpler than minor-closed classes, hereditary graph classes include a wider range of important classes, including planar, complete, bipartite, bounded degree, perfect, permutation and interval graphs, and forests.

A class of graphs $\mathcal{C}$ is said to have *unbounded tree-width* if for any $n \in \mathbb{N}$ there is a graph $G \in \mathcal{C}$ such that $tw(G) > n$. Similarly for the terms *unbounded clique-width* or *unbounded rank-width*.

Courcelle and Olariu [21] showed that bounded tree-width always implies bounded clique-width with the following result:

Theorem 1.4 ([21])**.** *For every graph* G, $cw(G) \leqslant 2^{tw(G)+1} + 1$.

However, bounded clique-width does not always imply bounded tree-width. For example, the class of complete graphs has bounded clique-width but unbounded tree-width.

Clearly from the definition, $lcw(G) \geqslant cw(G)$. Hence, a graph class of unbounded clique-width is also a class of unbounded linear clique-width. Likewise, a class with bounded linear clique-width is also a class of bounded clique-width. The two are frequently both bounded or both unbounded. When proving a class has bounded clique-width we usually create a bounded linear clique-width expression. However, this is not always the case – cographs are an example of a hereditary graph class that has bounded clique-width but unbounded linear clique-width [15, 37].

Useful results establishing the relationship of the clique-width of a graph and related graphs are as follows:

Lemma 1.5 (Gurski [35])**.** *If* $G = (V, E)$ *is a graph and* $v \in V$ *then*

$$\frac{1}{2} cw(G) \leqslant cw(G - v) \leqslant cw(G).$$

In particular, if H is an induced subgraph of G then the clique-width of H is at most the clique-width of G, so in this respect, clique-width is more compatible with hereditary classes of graphs than tree-width.

Lemma 1.6 (Courcelle and Olariu [21])**.** *Let* G *be a graph then*

$$\frac{1}{2} cw(G) \leqslant cw(\overline{G}) \leqslant 2\, cw(G)$$

where $\overline{G}$ *is the complement of* G*.*

Lemma 1.7 (Lozin and Rautenbach [46])**.** *Let* G *be a bipartite graph then*

$$\frac{1}{4} cw(G) \leqslant cw(\widetilde{G}) \leqslant 4\, cw(G)$$

where $\widetilde{G}$ *is the bipartite complement of* G*.*

Lemma 1.8 (Gurski and Wanke [38])**.** *Let* G *be a graph then*

$$\frac{(tw(G) + 1)}{4} \leqslant cw(L(G)) \leqslant 2\, tw(G) + 2$$

where $L(G)$ *is the line graph of of* G.

A useful survey of the status of knowledge regarding clique-width in hereditary graph classes at the time of commencing this project can be found in [23].

1.3.3 Rank-width

A related parameter is that of rank-width which was introduced by Oum and Seymour in 2006 [54].

Let $G = (V, E)$ be a graph and X, Y be disjoint subsets of V. Let M be the adjacency matrix of G over $GF(2)$. The adjacency matrix of a *cut* (X, Y) is the $0-1$ matrix $M[X, Y]$ whose rows are indexed by X, columns are indexed by Y, and the entry in row i and column j is equal to 1 if and only if i and j are connected by an edge of G. We define the *rank* of (X, Y), $\mathrm{rk}_G(X, Y)$, as $\mathrm{rk}(M[X, Y])$. The *cut-rank*, $\mathrm{cutrk}_G(X)$ of $X \subseteq V$, is defined by

$$\mathrm{cutrk}_G(X) = \mathrm{rk}_G(X, V \setminus X).$$

A *subcubic tree* is a tree such that every vertex has exactly one or three incident edges. The pair $(T, \mathcal{L})$ is called a *rank-decomposition* of G if T is a subcubic tree and $\mathcal{L}$ is a bijection from V to the set of leaves of T. For an edge e of T, the two connected components of $T \setminus e$ induce a partition (A, B) of the set of leaves of T. Letting $X = \mathcal{L}^{-1}(A)$ we define:

Definition 1.9. The *width* of edge e is $\mathrm{cutrk}_G(X)$. The *width* of $(T, \mathcal{L})$ is the maximum width of all edges of T. The *rank-width* of G, denoted $rw(G)$, is the minimum width of all rank-decompositions of G. (If $|V(G)| \leqslant 1$, we define $rw(G) = 0$.)

The relationship between the clique-width and rank-width of a graph is given by:

Lemma 1.10 ([54])**.** *For any simple graph* G

$$rw(G) \leqslant cw(G) \leqslant 2^{rw(G)+1} - 1$$

Hence, a graph class has bounded clique-width if and only if it has bounded rank-width.

Let v be a vertex of a simple graph G. The operation which replaces the subgraph of G induced by $N(v)$ by its complement is called *local complementation* and is denoted $G * v$. The operation of removing vertex v from G is denoted $G - v$. A graph H is *locally equivalent* to G if H can be obtained from G by a sequence of local complementations, and is a *vertex-minor* of G if it can be obtained from G by a sequence of local complementations and vertex removals.

We will use the shorthand $G' = G * L$ where $L = (v_1, v_2, \dots)$ to signify that $G' = G * v_1 * v_2 * \dots$ i.e. the graph formed by a sequence of local complementations on the vertices $v_1, v_2, \dots$.

Rank-width is usually associated with vertex-minor closed classes for the following reason:

Lemma 1.11 (Oum [53]). *If* H *is locally equivalent to* G*, then the rank-width of* H *is equal to the rank-width of* G*. If* H *is a vertex-minor of* G*, then the rank-width of* H *is at most the rank-width of* G*.*

For an edge uv of G, the *pivot* on uv is the graph $G * u * v * u$. That this is well-defined (i.e. u and v can be interchanged in the definition) is established in [53], and the effect of this process is to complement the edges between the three sets of vertices $N(u) \cap N(v)$, $N(u) \setminus N(v)$ and $N(v) \setminus N(u)$.

A *circle graph* is an intersection graph of finitely many chords of a circle and the class of circle graphs is vertex-minor closed. A major recent result analogous to the Grid Minor Theorem but relating to vertex-minor-closed classes came from Geelen, Kwon, McCarty and Wollan [33, 51] who established that boundedness of rank-width in such classes follows if and only if the class excludes a circle graph as a vertex-minor. In effect, this result means that the class of circle graphs (the intersection graphs of sets of chords of a circle) plays a similar role in vertex-minor-closed classes as planar graphs do for minor-closed classes.

1.4 Graph parameters and the evaluation problem

Courcelle [18] in 1990 and later Courcelle, Makowsky and Rotics [20] in 2000 established the relationship of the evaluation problem to tree-width and clique-width. They considered evaluation problems expressible using a particular type of logic known as monadic second order (MSO) logic, two major variants of which are MSO_1 (where only vertex and vertex set variables are allowed), and MSO_2 (where edge and edge set variables are also allowed).

Theorem 1.12 ([18]). *For every evaluation problem definable in* MSO_2 *logic there exists a function* f *such that for any* n*-vertex graph* G *with* $tw(G) < k$ *the problem can be solved in time at most* $f(k)n$*.*

Theorem 1.13 ([20]). *For every evaluation problem definable in MSO_1 logic there exists a function f such that for any n-vertex graph G with $cw(G) < k$ the problem can be solved in time at most $f(k)n$.*

MSO_2 logic has greater expressive power than MSO_1. For example, the existence of a perfect matching or of a Hamiltonian cycle in a graph is expressible in MSO_2, but not in MSO_1. On the other hand, there are more graph classes with bounded clique-width than bounded tree-width. MSO_1 logic still covers important problems such as maximum independent set, minimum dominating set, and k-colourability (Also note that it has been shown that the Hamilton cycle problem can still be solved in polynomial time for graphs of bounded clique-width [9]).

1.5 Sparsity and Density

The least number of forests that can cover the edges of a graph is called its *arboricity*. A graph with arboricity bounded by $k \in \mathbb{N}$ is called k-*uniformly sparse* or just *sparse*. A graph class is k-uniformly sparse if it does not contain a graph of arboricity greater than k.

The following theorems are useful:

Theorem 1.14 (Nash-Williams, 1964 [52]). *The edges of a graph $G = (V, E)$ can be covered by at most k forests if and only if $\|G[U]\| \leqslant k(|U| - 1)$ for every non-empty set $U \subseteq V$, where $\|G[U]\|$ denotes the number of edges in the subgraph induced by the vertices U.*

Theorem 1.15 (Kostochka, 1982 [30]). *There exists a constant $c \in \mathbb{R}$ such that, for every $t \in \mathbb{N}$, every graph G of average degree $d(G) \geqslant ct\sqrt{\log t}$ contains K_t as a minor.*

A *dense* hereditary graph class is one that is not k-uniformly sparse for some $k \in \mathbb{N}$, or in other words, by Theorem 1.14, for any k there is a graph in the class that has average degree greater than k. By Theorem 1.15, for any $t \in \mathbb{N}$ we can set $k \geqslant ct\sqrt{\log t}$ so that the class contains a graph with a K_t minor.

It is easy to show that $tw(K_t) = t - 1$ for all positive integers $t \geqslant 2$. Theorems 1.14 and 1.15 tell us that all dense hereditary graph classes contain a graph with a K_t minor for any positive integer $t \geqslant 2$ and so by Lemma 1.2 have unbounded tree-width.

The significance of sparsity to clique-width comes from the following. Combining results from [21] with results from Gurski and Wanke [36] gives us certain graph classes for which tree-width and clique-width are either both bounded or both unbounded:

Theorem 1.16 ([21, 36])**.** *If $\mathcal{C}$ is a collection of graphs such that every graph $G \in \mathcal{C}$ either*

(i) has maximum vertex degree bounded by a constant, or

(ii) excludes a fixed graph H as a minor, or

(iii) has bounded arboricity,

then $\mathcal{C}$ has unbounded clique-width if and only if it has unbounded tree-width.

We consider the existence of obstructions that prevent a hereditary graph class from having bounded clique-width separately for sparse and dense hereditary classes.

1.6 Well-quasi-ordering in hereditary graph classes

Well-quasi-ordering in a graph class is a particularly attractive property, as explained below, and therefore attracts much attention.

A reflexive and transitive relation is called a *quasi-ordering*. A quasi-ordering $\leqslant$ on a set X is a *well-quasi-ordering* if for every infinite sequence $x_0, x_1, \dots$ in X there are indices $i < j$ such that $x_i \leqslant x_j$. An *infinite antichain* is an infinite sequence $x_0, x_1, \dots$ in X such that there is no pair of indices i, j such that $x_i \leqslant x_j$ or $x_j \leqslant x_i$.

It can be shown that a quasi-ordering $\leqslant$ on X is a well-quasi-ordering if and only if X contains neither an infinite antichain nor an infinitely strictly decreasing sequence $x_0 > x_1 > \dots$ (see [30] page 348). Given the fact that no strictly descending sequence of induced subgraphs (or minors) can be infinite, then if our relationship is 'is an induced subgraph of' (or 'minor of') being well-quasi-ordered is equivalent to the non-existence of an infinite antichain.

Well-quasi-ordering is central to the Graph Minor Theorem [58], that tells us that the finite graphs are well-quasi-ordered by the minor relation. Consequently, any minor-closed graph class, such as planar graphs, is finitely defined under the minor relationship – that is, such a class can be defined by a finite number of minimal forbidden minors. Most famously, Wagner's Theorem states that a finite graph is planar if and only if it contains neither K_5 nor $K_{3,3}$ as a minor [60]. We say a graph G is H-*minor-free* if it does not contain a minor of the graph H, and a class $\mathcal{C}$ has an *excluded minor* H if every graph in the class is H-minor-free.

Hereditary graph classes – that is, classes closed under taking induced subgraphs – are not, in general, well-quasi-ordered due to the presence of infinite antichains under the induced subgraph relationship. An antichain under the induced subgraph relationship is a set $\mathcal{A}$ of graphs such that if G and H are distinct graphs in $\mathcal{A}$ then neither $H \leqslant G$ nor $G \leqslant H$ (Here, '$\leqslant$' means 'is an induced subgraph of').

Daligault, Rao and Thomassé [27] questioned whether well-quasi-ordering by induced subgraphs always implies bounded clique-width for hereditary classes. This was answered in the negative by Lozin, Razgon and Zamaraev [49] with the example of 'power graphs'. Dabrowski, Lozin and Paulusma [24] subsequently showed that in (K_3, H)-free classes (that is, classes with two minimal forbidden induced subgraphs where one of them is K_3 and the other an arbitrary graph H) well-quasi-ordering by induced subgraphs does imply bounded clique-width.

A hereditary class of graphs $\mathcal{C}$ can be defined by a unique set of minimal forbidden graphs $\{H_1, H_2, \dots\}$ i.e., every graph $G \in \mathcal{C}$ is H_i-free for $i = 1, 2, \dots$. Such a set of minimal forbidden graphs must necessarily be an antichain, but may be finite or infinite. Classes that have a finite number of minimal forbidden graphs are attractive since this finite set is a convenient 'certificate' to use in an algorithm to easily check if an input graph is in the class. Unsurprisingly, much recent research into obstructions to bounded tree-width and clique-width in hereditary classes has focussed on *finitely defined* classes (i.e., the list of minimal forbidden graphs is finite) – see [3, 5, 22, 25, 26, 48]. However, important though this is, it does not give the complete picture.

1.7 Minimal classes or antichains of unbounded clique-width

Having established the importance of bounded clique-width in the study of evaluation problems defined on a graph, we might next ask: What are the obstructions that prevent a hereditary graph class from having bounded clique-width?

Minimal hereditary graph classes of unbounded clique-width are one such obstruction. A hereditary class $\mathcal{C}$ is minimal of unbounded clique-width (hereafter shortened to just *minimal*) if every proper hereditary subclass $\mathcal{D}$ has bounded clique-width. In other words, a hereditary graph class $\mathcal{C}$ is minimal if, for any proper hereditary subclass $\mathcal{D}$ formed by adding just one more forbidden graph, $\mathcal{D}$ has bounded clique-width.

A wider concept is that of *limit class*. A hereditary class $\mathcal{C}$ is a limit class if there exists a sequence of distinct hereditary classes $\mathcal{C}_1 \supset \mathcal{C}_2 \supset \dots$ each of unbounded clique-width such that $\mathcal{C} = \cap_i \mathcal{C}_i$. A minimal limit class is called a *boundary class*. The idea of boundary classes was introduced by Alekseev [7] and was later used by Lozin and Milanic in [45] in the context of obstructions to bounded clique-width. A boundary class must be contained in every hereditary graph class of unbounded clique-width.

Whilst minimal classes are boundary classes, there are boundary classes that are not minimal.

Lozin and Rautenbach identified two such (non-minimal) boundary classes, tripods (a forest in which every connected component has at most 3 leaves) denoted $\mathcal{S}$ and line graphs of tripods $\mathcal{T}$, and proved the following result:

Theorem 1.17 ([43, 47]). *If $\mathcal{C}$ is a finitely defined hereditary class that contains $\mathcal{S}$ or $\mathcal{T}$, then the tree-width and clique-width of graphs in $\mathcal{C}$ is unbounded.*

More recently, Lozin and Razgon extended this result to a characterization of hereditary classes of unbounded tree-width that are finitely defined:

Theorem 1.18 ([48]). *The tree-width of graphs in a hereditary class defined by a finite set $\mathcal{F}$ of forbidden induced subgraphs is bounded if and only if $\mathcal{F}$ includes a complete graph, a complete bipartite graph, a tripod and the line graph of a tripod.*

Boundary classes have proved useful in studying finitely defined hereditary graph classes but are not the only way to characterize obstructions to bounded clique-width.

An alternative characterization is given by Dawar and Sankaran [29] who have shown that a hereditary graph class of unbounded clique-width must contain at least one of the following:

(a) A minimal class, or

(b) An *antichain of unbounded clique-width*, that is, a sequence of graphs $A_1, A_2, \ldots$ such that there is no pair of indices i, j such that $A_i \leqslant A_j$ or $A_j \leqslant A_i$ and $cw(A_i) \to \infty$ as $i \to \infty$.

In particular, they show that:

Theorem 1.19 ([29]). *A minimal class cannot contain an antichain of unbounded clique-width.*

A class is *atomic* if it cannot be expressed as the union of two proper subclasses.

Lemma 1.20. *Every minimal class is atomic.*

Proof. Suppose $\mathcal{C}$ is a hereditary class with unbounded clique-width but is not atomic, so that $\mathcal{C} = \mathcal{D} \cup \mathcal{E}$ where $\mathcal{D}$ and $\mathcal{E}$ are two proper subclasses. Since $\mathcal{C}$ has unbounded clique-width then either $\mathcal{D}$ or $\mathcal{E}$ (or both) have unbounded clique-width. Without loss of generality let $\mathcal{D}$ have unbounded clique-width, then $\mathcal{D} \subsetneq \mathcal{C}$ so that $\mathcal{C}$ is not minimal. □

We use the following version of Fraïssé's Theorem expressed in terms of hereditary graph classes:

Theorem 1.21 (Fraïssé's [32]; see also Hodges [40], Section 7.1)**.** *For a hereditary class of graphs $\mathcal{C}$ the following are equivalent:*

1. *$\mathcal{C}$ is atomic,*
2. *$\mathcal{C}$ satisfies the joint embedding property (that is, given any two graphs D and E in $\mathcal{C}$ we can find a third graph $F \in \mathcal{C}$ containing induced subgraphs isomorphic to D and E),*
3. *$\mathcal{C}$ is the class of the induced subgraphs of a single infinite connected graph G.*
4. *there exists a sequence of graphs $G_1 \leqslant G_2 \leqslant \ldots$ in $\mathcal{C}$ such that for all $H \in \mathcal{C}$ there exists i such that H is an induced subgraph of G_i.*

Corollary 1.22. *A minimal class can always be represented as the set of induced subgraphs of a single infinite connected graph.*

This gives us a starting point in the search for minimal classes since we need only consider single infinite connected graphs.

1.8 Proving a hereditary class is minimal

There are two steps to proving a hereditary class is minimal – proving that the class has unbounded clique-width and then proving it is minimal. Proving a class $\mathcal{C}$ has unbounded clique-width usually follows one of the following routes:

1. From the definition, showing that $\mathcal{C}$ contains a series of graphs G_n such that $cw(G_n) \to \infty$ as $n \to \infty$. Finding the exact cw for a graph can be hard and therefore we usually find a lower bound $cw(G_n) \geqslant l_n$ such that $l_n \to \infty$ as $n \to \infty$. This method is used in Chapters 2 and 6 and papers [12, 26, 34, 42, 49, 50].
2. For sparse classes, using the relationship with tree-width and showing that $\mathcal{C}$ contains a series of graphs G_n each containing a minor H_n such that $cw(H_n) \to \infty$ as $n \to \infty$ and observing that $cw(G_n) \geqslant cw(H_n)$. This method is used in Chapters 4 and 5.
3. For dense classes, using the relationship with rank-width and showing that $\mathcal{C}$ contains a series of graphs G_n containing a vertex-minor H_n such that $cw(H_n) \to \infty$ as $n \to \infty$ and observing that $cw(G_n) \geqslant cw(H_n)$. This method is used in [8, 13, 17].

4. using MSO (monadic second order logic) transductions, showing that the class 'interprets' arbitrarily large grids. This method is not used here – for an explanation see [29].

Proving a class is minimal requires showing that any proper subclass $\mathcal{D}$ has bounded (linear) clique-width. An arbitrary graph $G \in \mathcal{D}$ is decomposed into sections or 'panels' (path or grid components) and a clique-width expression created so that each section requires a bounded number of labels to construct and the sections can be joined together also with a bounded number of labels.

In [13, 44] this is done using Menger's theorem, in Chapter 3 we use a new method of 'veins' and 'slices' (inspired by a corresponding structure in the study of permutations [4]) and in Chapter 5 and [29] an observation regarding the linear clique-width of a bounded interval in the class.

We can also create further minimal classes by taking complements, bipartite complements or (possibly) line graphs of known minimal classes. We define:

(i) $\overline{\mathcal{C}_1}$ to be the *complement class* of a hereditary class of graphs $\mathcal{C}_1$ where $\overline{G} \in \overline{\mathcal{C}_1}$ if and only if $G \in \mathcal{C}_1$.

(ii) $\widetilde{\mathcal{C}_2}$ to be the *bipartite complement class* of a hereditary class of bipartite graphs $\mathcal{C}_2$ where $\widetilde{G} \in \widetilde{\mathcal{C}_2}$ if and only if $G \in \mathcal{C}_2$.

(iii) $L(\mathcal{C}_3)$ to be the *line graph class* derived from a hereditary class $\mathcal{C}_3$ where $G_L \in L(\mathcal{C}_3)$ if and only if $G_L \leqslant L(G)$ for some $G \in \mathcal{C}_3$.

First, notice that each of these classes is hereditary. Furthermore:

Lemma 1.23. *$\overline{\mathcal{C}_1}$ has unbounded clique-width if and only if $\mathcal{C}_1$ does. $\widetilde{\mathcal{C}_2}$ has unbounded clique-width if and only if $\mathcal{C}_2$ does. If $\mathcal{C}_3$ is a sparse class, then $L(\mathcal{C}_3)$ has unbounded clique-width if and only if $\mathcal{C}_3$ does.*

Proof. This follows from Lemmas 1.6, 1.7 and 1.8 respectively. Note we restrict this to sparse $\mathcal{C}_3$ since we are using the fact that tree-width and clique-width are either both bounded or both unbounded in a sparse class (Theorem 1.16). If $\mathcal{C}_3$ is dense then the clique-width of $L(\mathcal{C}_3)$ is always unbounded. □

Lemma 1.24. *(i) $\overline{\mathcal{C}_1}$ is a minimal class (antichain of unbounded clique-width) if and only if $\mathcal{C}_1$ is a minimal class (antichain of unbounded clique-width).*

(ii) $\widetilde{\mathcal{C}_2}$ is a minimal class (antichain of unbounded clique-width) if and only if $\mathcal{C}_2$ is a minimal class (antichain of unbounded clique-width).

Proof. Suppose $\mathcal{C}_1$ is a minimal class. If $H, G \in \mathcal{C}_1$ then $H \leqslant G$ if and only if $\overline{H} \leqslant \overline{G}$. Since $\mathcal{C}_1$ does not contain an antichain of unbounded clique-width it follows that $\overline{\mathcal{C}_1}$ does not either. Thus $\mathcal{C}_1$ must contain a minimal class. Furthermore, forbidding one graph from $\mathcal{C}_1$ results in forbidding one graph in $\overline{\mathcal{C}_1}$, so $\overline{\mathcal{C}_1}$ must also be minimal.

The same approach applies with $\mathcal{C}_2$ and $\widetilde{\mathcal{C}_2}$. □

The position is more complicated for a line graph class. If we consider the graphs H and G in Figure 1.1, we have $H \not\leqslant G$ but $L(H) \leqslant L(G)$. This suggests that it may be possible the class C_3 contains an antichain of unbounded clique-width whilst $L(C_3)$ does not, and therefore contains a minimal class. This remains an open question:

Conjecture 1.25. *There exists a hereditary class $\mathcal{C}$ of unbounded clique-width that does not contain a minimal class whose line graph class $L(\mathcal{C})$ contains a minimal class.*

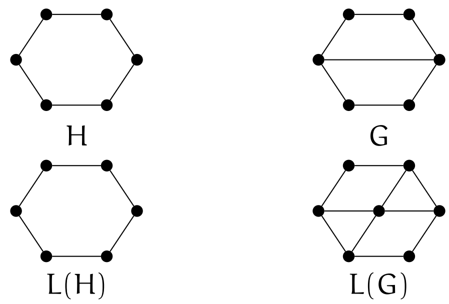


Figure 1.1: Example: $H \not\leqslant G$ but $L(H) \leqslant L(G)$

1.9 Clique-width in dense hereditary classes – First Framework

Following the introduction of clique-width as a graph parameter, a number of bounded clique-width classes (e.g. cliques, bicliques, distance-hereditary graphs [34]) and unbounded clique-width classes (e.g. unit interval graphs, permutation graphs [34], bipartite permutation graphs [12] and split graphs [50]) were identified. It was also shown that every superfactorial graph class has unbounded clique-width [10] (a class X is said to be *superfactorial* if the number X_n of n-vertex graphs in X satisfies $X_n > n^{cn}$ for some constant $c > 0$).

Following speculation regarding the existence of minimal classes of unbounded clique-width, in 2011, Lozin [44] demonstrated the first such minimal classes - bipartite permutation graphs and unit interval graphs. More recently further classes have been identified, in Atminas, Brignall, Lozin and Stacho [8] (bichain graphs and split permutation graphs), Collins, Foniok, Korpelainen, Lozin and Zamaraev [17] (a grid structure with a countably infinite number of minimal classes) and Dawar and Sankaran [29] (power graphs).

The core of this thesis is contained in Chapters 2 and 3 where we create a framework for dense minimal hereditary graph classes of unbounded clique-width. The framework brings together all but one (the power graphs) of the previously discovered minimal classes into a single consistent framework. The framework consists of hereditary graph classes constructed by taking the finite induced subgraphs of an infinite graph $\mathcal{P}^\delta$ whose vertices form a two-dimensional array and whose edges are defined by three objects, collectively denoted as a triple $\delta = (\alpha, \beta, \gamma)$. The hereditary graph class $\mathcal{G}^\delta$ is the set of all finite induced subgraphs of $\mathcal{P}^\delta$. Though we defer full definitions until Chapter 2, the components of the triple define edges between consecutive columns (α), between non-consecutive columns (β 'bonds'), and within columns (γ) as follows:

(a) α is an infinite word from the alphabet $\{0, 1, 2, 3\}$. The four types of α-edge sets between consecutive columns can be described as a matching (0), the complement of a matching (1), a half-graph (2) and the complement of a half-graph (3) (for half-graph definition see Section 1.2.1).

(b) β is a symmetric subset of pairs of natural numbers (x, y) where $|x - y| > 1$. If $(x, y) \in \beta$ then every vertex in column x is adjacent to every vertex in column y.

(c) γ is an infinite binary word. If the j-th letter of γ is 0 then vertices in column j form an independent set and if it is 1 they form a clique.

An example of a graph from the framework is shown in Figure 1.2.

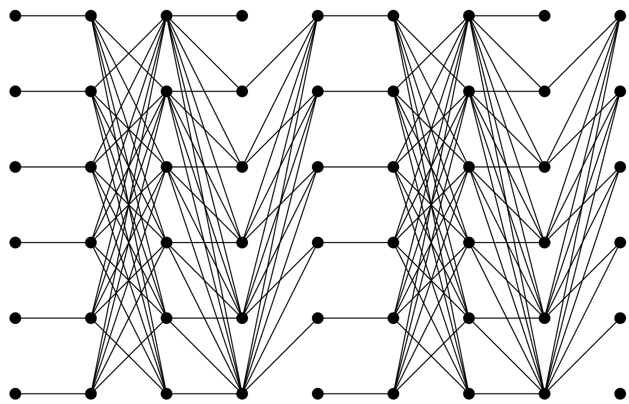


Figure 1.2: The graph $H^\delta_{1,1}(9,6)$ from the framework, where $\delta = (\alpha, \beta, \gamma)$ with $\alpha = 01230123\cdots$, $\beta = \emptyset$ and $\gamma = 0^\infty$

We show that these hereditary graph classes $\mathcal{G}^\delta$ have unbounded clique-width if and only if a parameter $\mathcal{N}^\delta$ measuring the number of distinct neighbourhoods between any two rows of the grid, is unbounded – see Theorem 2.17. We denote Δ as the set of δ-triples for which $\mathcal{G}^\delta$ has unbounded clique-width.

Furthermore, we define a subset $\Delta_{min} \subset \Delta$ such that if $\delta \in \Delta_{min}$ the hereditary graph class $\mathcal{G}^\delta$ is minimal both of unbounded clique-width and of unbounded *linear* clique-width (Definitions in Section 1.3.2 and result Theorem 3.11). Referring to $\delta^* = \delta_{[a,a+b]}$ as a *factor of* δ being a subset of δ defining all edges between vertices in columns $a, a+1, \dots, a+b$, these 'minimal' δ-triples are characterised by:

(a) $\delta \in \Delta$,

(b) δ is $\mathcal{N}^\delta$*-bounded recurrent* (i.e. any factor δ^* of δ repeats an infinite number of times, and the subgraphs induced on the columns between two consecutive disjoint copies of δ^* (the δ-factor 'gap') have bounded $\mathcal{N}^\delta$ (always true for almost periodic δ)), and

(c) a bound on a parameter $\mathcal{M}^\beta$ defined by the bond set β, which is a measure of the number of distinct neighbourhoods between intervals of a single row.

All hereditary graph classes previously shown to be minimal of unbounded clique-width (except the power graphs) fit this grid framework i.e. they are defined by a δ-triple in Δ_{min}. This is demonstrated in Table 1.1 which shows their corresponding $\delta = (\alpha, \beta, \gamma)$ values from the framework. The power graphs are considered in Chapter 5.

Name	α	β $(x, y \in \mathbb{N})$	γ
Bipartite permutation [44]	2^∞	$\emptyset$	0^∞
Unit interval [44]	2^∞	$\emptyset$	1^∞
Bichain [8]	$(23)^\infty$	$(2x, 2x+2y+1)$	0^∞
Split permutation [8]	$(23)^\infty$	$(2x, y) : y > 2x+1$	$(01)^\infty$
$\alpha \in \{0,1\}$ [17]	periodic	$\emptyset$	0^∞

Table 1.1: Hereditary graph classes previously proven to be minimal

The viewpoint provided by the framework offers a fuller understanding of the landscape of (the uncountably many known) minimal hereditary classes of unbounded clique-width. This landscape is in stark contrast to minor-closed classes with respect to tree-width, where planar graphs are the unique minimal minor-closed class of graphs of unbounded treewidth (see Section 1.3.1), or to vertex-minor closed classes with respect to rank-width, where circle graphs are the unique minimal vertex-minor-closed class of unbounded rank-width (see Section 1.3.3).

As it happens, any proper subclass of a minimal class from our framework also has bounded linear clique-width. However, beyond our framework there do exist classes that have bounded clique-width but unbounded linear clique-width, see [6] and [15].

In Chapter 2 we set out exactly how the framework of $\mathcal{G}^\delta$ hereditary graph classes is constructed, followed by the proof determining which hereditary classes $\mathcal{G}^\delta$ have unbounded clique-width. Proving a class has unbounded clique-width is done from first principles, using a new method, by identifying a lower bound for the number of labels required for a clique-width expression for an $n \times n$ square graph, using distinguished coloured vertex sets and showing such sets always exist for big enough n using Ramsey theory. For those classes which have bounded clique-width, we prove this by providing a general clique-width expression for any graph in the class, using a bounded number of labels.

In Chapter 3 we prove that the class $\mathcal{G}^\delta$ is minimal of unbounded clique-width if $\delta \in \Delta_{min}$. To do this we introduce an entirely new method of 'veins and slices', partitioning the vertices of an arbitrary graph in a proper subclass of $\mathcal{G}^\delta$ into sections we call 'panels' using vertex colouring. We then create a recursive linear clique-width expression to construct these panels in sequence, allowing recycling of labels each time a new panel is constructed, so that an arbitrary graph can be constructed with a bounded number of labels.

Previous papers on minimal hereditary graph classes of unbounded clique-width have focused mainly on bipartite graphs. The introduction of β-bonds and γ-cliques significantly broadens the scope of proven minimal classes.

We show that in this framework there are uncountably many minimal hereditary graph classes of unbounded clique-width (Theorem 3.14).

In [17] it was conjectured that, in a simpler version of this framework equivalent to classes where $\delta = (\alpha, \emptyset, 0^\infty)$, such a class was minimal if and only if α is non-zero and almost periodic. In Section 3.4 we demonstrate a counterexample of a minimal class defined by a recurrent but not almost periodic word.

In Section 3.5 we provide some examples of new hereditary graph classes that are minimal of unbounded clique-width revealed by this approach.

1.10 Clique-width in sparse hereditary classes

From Theorem 1.16 a sparse hereditary class has bounded clique-width if and only if it has bounded tree-width, therefore the study of sparse classes tends to focus on the

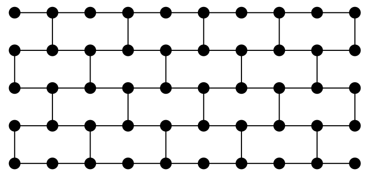


Figure 1.3: The 4×4 wall $W_{4\times 4}$

more extensively researched tree-width. The question becomes: What are the obstructions that prevent a sparse graph class from having bounded tree-width? We start by defining a *wall*.

An $(m \times n)$-*wall* is a graph whose edges are visually equivalent to the mortar lines of a stretcher-bonded clay brick wall with m rows of bricks each of which is n bricks long. More precisely, we can define the wall $W_{m\times n} = (V, E)$ using a square grid of the usual (x, y) Cartesian coordinates.

$$\begin{aligned}
V &= \{(x,y) : 0 \leqslant x \leqslant 2n+1, 0 \leqslant y \leqslant m\} \\
E_H &= \{(x,y)(x+1,y) : (x,y),(x+1,y) \in V\} \\
E_V &= \{(x,y)(x,y+1) : (x,y),(x,y+1) \in V, x+y = 0 (\text{mod } 2)\} \\
E &= E_H \cup E_V.
\end{aligned}$$

See example of $W_{4\times 4}$ in Figure 1.3.

It is well-known that the following t-*basic obstructions* are obstructions to bounded tree-width: for arbitrarily large t, (1) the complete graph K_t, (2) the complete bipartite graph $K_{t,t}$, (3) a subdivision of the $(t \times t)$-wall and (4) the line graph of a subdivision of the $(t \times t)$-wall.

Although tree-width was originally associated with minor-closed graph classes, recent interest in tree-width has focussed on hereditary graph classes, sparked by a paper by Aboulker, Adler, Kim, Sintiari and Trotignon containing the result:

Theorem 1.26 (Induced Grid Theorem for Minor-Free Graphs [1])**.** *For every graph H there is a function $f_H : \mathbb{N} \to \mathbb{N}$ such that every H-minor-free graph of tree-width at least $f_H(t)$ contains an induced subgraph isomorphic to a subdivision of a $(t \times t)$-wall or the line graph of a subdivision of a $(t \times t)$-wall.*

This was followed by a result from Korhonen:

Theorem 1.27 (Corollary to The Grid Induced Minor Theorem [41])**.** *For every $\Delta \in \mathbb{N}$ there is a function $f_\Delta : \mathbb{N} \to \mathbb{N}$ such that every graph with degree at most Δ and tree-width at least $f_\Delta(t)$ contains an induced subgraph isomorphic to a subdivision of a $(t \times t)$-wall or the line graph of a subdivision of a $(t \times t)$-wall.*

Thus, a hereditary graph class with an excluded minor or of bounded vertex degree has unbounded tree-width if and only if it contains arbitrarily large subdivisions of a wall or the line graph of a subdivision of a wall. This suggests the possibility of a similar characterization of the larger family of uniformly sparse graphs.

Tree-width in hereditary graph classes is currently a very active area of research. In particular, Abrishami, Alecu, Chudnovsky, Dibek, Hajebi, Rzążewski, Spirkl and Vušković have contributed to a series of papers on the topic of induced subgraphs and tree decompositions, e.g. [2, 3, 5]. This series is particularly interested in graph classes that are t-*clean* meaning that graphs with large tree-width contain a large t-basic obstruction.

However, the four named objects are not the only obstructions to bounded tree-width in sparse hereditary classes. Counterexamples have recently been found by Sintiari and Trotignon [59], Davies [28] and Bonamy, Bonnet, Déprés, Esperet, Geniet, Hilaire, Thomassé and Wesolek [11].

Definition 1.28. A hereditary graph class is KKW-*free* if there exists $t \in \mathbb{N}$ such that the class does not contain the complete graph K_t, the complete bipartite graph $K_{t,t}$, a subdivision of the $(t \times t)$-wall or the line graph of a subdivision of the $(t \times t)$-wall.

Our interest is in KKW-free classes that nevertheless have unbounded tree-width, that is, that are *not* t-clean.

In sparse classes, the simplest structure to consider is a forest, but, trivially, the tree-width of a forest is bounded. The natural place to start looking for obstructions to bounded tree-width is, therefore, graphs of arboricity two. Using Ramsey theory [30, Proposition 9.4.1] we know that for every $r \in \mathbb{N}$ there is an $n \in \mathbb{N}$ such that every connected graph of order at least n contains a clique, a star or a path of order r as an induced subgraph. Since large trees contain long induced paths or big induced stars, this suggests that in classes of arboricity two the natural structures to consider are 'path-path' (two forests of paths), 'path-star'(a forest of paths and a forest of stars) and 'star-star' (two forests of stars). Given that walls and line graphs of walls are the only 'path-path' obstructions to bounded tree-width as a consequence of Theorem 1.27, we seek further obstructions to bounded tree-width in another family of classes of arboricity two, namely 'path-star' classes.

Definition 1.29. A finite graph (V, E) is a *path-star* graph if its edges E can be partitioned into two sets, E_P and E_S, so that

1. (V, E_P) is a forest of paths, and
2. (V, E_S) is a forest of stars.

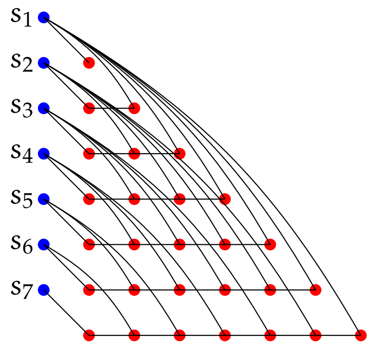


Figure 1.4: A 7-sail

In Chapter 4 we consider a particular type of hereditary *path-star class* being the collection of all the finite induced subgraphs of a single infinite graph whose edges can be partitioned into a path (or forest of paths) together with a forest of stars, where the leaves of the stars (but not the internal vertices of the stars) may embed in the paths. We denote a path-star class $\mathcal{R}^\alpha$ where α is an infinite word over alphabet $\mathbb{N}$ (see formal Definition 4.10). Our study of path-star graph classes has led to the discovery of a family of objects, t-*sails*, with tree-width at least $t-1$ (Lemma 4.11).

Definition 1.30. A path-star graph $(V, E_P \cup E_S)$ is a t-sail if there exists a set $V_S = \{s_1, \ldots, s_t\} \subseteq V$ so that

1. No edge in E_P is incident with a vertex in V_S,
2. The graph $(V \setminus V_S, E_P)$ comprises t components $P_1, \ldots, P_t$ (all paths), and
3. For all $1 \leqslant i \leqslant j \leqslant t$ there exists $v \in P_i$ such that $s_j v \in E_S$.

An example is shown in Figure 1.4.

In fact, independently developed, t-sails are a generalisation of structures previously observed by Pohoata [55] and Davies [28]. In particular, in [28] a type of t-sail was used to disprove a conjecture that tree-width is always $O(\log|V(G)|)$ for a graph G in a hereditary class that is KKW-free.

In Chapter 4 we show that a path-star class, defined by an infinite word over an infinite alphabet where each letter in the alphabet appears an infinite number of times in the word, contains a t-sail for arbitrarily large t and therefore has unbounded tree-width (Theorem 4.12). To distinguish this work from that recently undertaken on circle graphs by Hickingbotham, Illingworth, Mohar and Wood [39], we show that no path-star class defined by a word where at least two letters alternate more than four times is a subclass of circle graphs (Theorem 4.13).

We identify a collection of *nested* words (Section 4.2) with a recursive structure that exhibit interesting characteristics when used to define a hereditary path-star graph class:

Theorem 4.16. If α is a nested word then the path-star class $\mathcal{R}^\alpha$ is KKW-free.

Theorem 4.22. If α is a nested word then for every $t \geqslant 1$ there is a positive integer valued function $f_\alpha(t)$ such that every graph in $\mathcal{R}^\alpha$ of tree-width at least $f_\alpha(t)$ contains a t-sail as an induced subgraph.

This shows that for a path-star class defined by a nested word, t-sails perform a similar role as walls do for minor excluded or bounded degree classes, in that they are the basic object obstructing bounded tree-width, albeit that in this case a t-sail is not one fixed graph like a wall but a more general construction.

In Section 1.6 we noted that hereditary class may be finitely or infinitely defined. We show that the following is a consequence of Theorem 1.18:

Theorem 4.23. A hereditary class of graphs of unbounded tree-width that is KKW-free is infinitely defined.

Given the uncountable number of dense minimal classes of unbounded clique-width revealed in Chapter 3, it is natural to ask whether there are any sparse minimal classes of unbounded tree-width (or clique-width). We show the following:

Theorem 4.26. If $\mathcal{C}$ is a hereditary class of graphs of bounded vertex degree or has an excluded minor then it does not contain a minimal class.

Theorem 4.28. If $\mathcal{R}^\alpha$ is a path-star hereditary class of graphs defined by a nested word α then it does not contain a minimal class.

Given these results we suggest:

Conjecture 1.31. *Sparse hereditary graph classes of unbounded tree-width do not contain a minimal class of unbounded tree-width.*

1.11 Clique-width in dense hereditary classes – Second Framework

In Chapter 5 we introduce an alternative framework of dense hereditary classes of unbounded clique-width, which are not covered by the framework in Chapter 2 (except to the extent that such a class might be present in the β-bonds which, by definition, are very general and could incorporate any hereditary class).

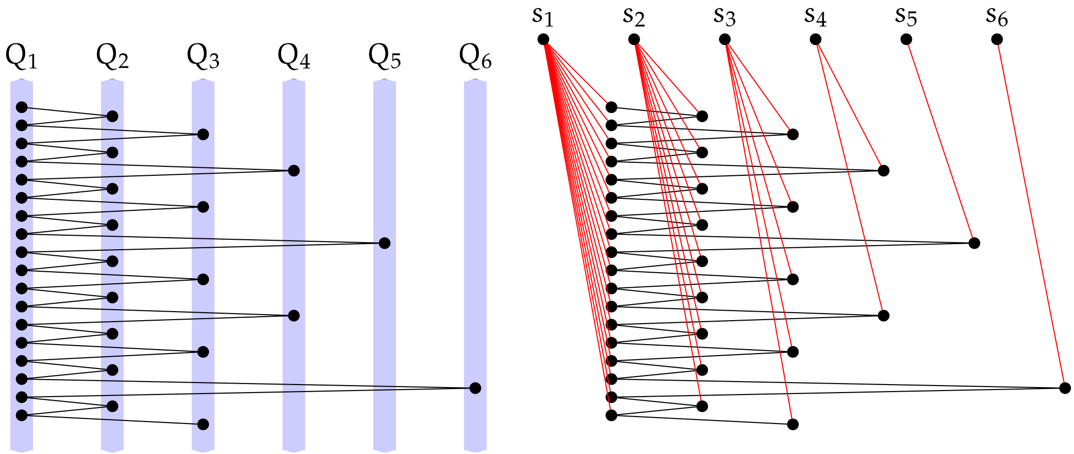


Figure 1.5: Comparing the path-clique graph $H^{(\alpha,\beta)}$ 'power graphs' (left) and the path-star graph R^{α} (right) where $\alpha = \kappa(2)$ and $\beta = \{(i,i) : i \in \mathbb{N}\}$ (path edges in black, clique edges in blue shaded areas, star-edges in red).

The graph operation of local complementation, where the subgraph induced by the neighbourhood of a vertex is replaced by its complement, can in some cases be used to turn a dense graph into a sparse one ('sparsification') and vice-versa ('densification'). Recall that the rank-width of a graph is unchanged by taking local complements (Lemma 1.11), that a class of graphs has unbounded rank-width if and only if it has unbounded clique-width (Lemma 1.10) and that, if the class is sparse, it has unbounded tree-width if and only if it has unbounded rank-width or clique-width (Lemma 1.16). Therefore, we can study the clique-width and rank-width of a dense class by sparsifying it to a sparse class and then considering its tree-width, a much better understood graph parameter.

The framework in Chapter 2 included all known minimal classes except for the 'power graphs' class (defined in Section 5.2.1 and shown in Figure 1.5) introduced by Lozin, Razgon and Zamaraev [49] and proven to be minimal by Dawar and Sankaran [29]. Notice that there is a structural similarity between the dense power graphs and the sparse path-star graphs introduced in Chapter 4, in the sense that they are both constructed as the induced subgraphs of an infinite graph built around an infinite path. Indeed, using only local complementation on one of the vertices in each clique together with some suitable vertex removal, it is quite easy to see that power graphs can be sparsified to graphs in the path-star class $\mathcal{R}^{\kappa(2)}$.

Using this idea of sparsifying a graph class by local complementation, in Chapter 5 we show that the power graphs belong to a wider family of dense graph classes with unbounded clique-width, constructed from an infinite path and an auxiliary graph of cliques and bicliques and that this family includes a collection of minimal classes.

In Section 5.2 we introduce *path-clique* graph classes, denoted $\mathcal{H}^{(\alpha,\beta)}$ and defined by a pair of objects (α, β). The class $\mathcal{H}^{(\alpha,\beta)}$ consists of the induced subgraphs of an infinite graph whose vertex set is indexed by the natural numbers. These vertices are partitioned into an infinite number of sets determined by α. Edges consist of the symmetric difference between the edge sets of two graphs, defined on the vertices:

1. an infinite path in vertex index order, and
2. a 'biclique graph' defined by β, where each set of the partition is a clique or independent set and such sets are complete or anticomplete to each other.

This family of graph classes is not covered by the framework in Chapter 2. Excluding β-bonds and cliques on columns, the framework in Chapter 2 only allows edges between consecutive columns. The sets in a path-clique graph class do not, in general, have this restriction.

We also introduce the following auxiliary graphs – the *class-path* P, the *biclique graph* $B^{(\alpha,\beta)}$ and the *star graph* $S^{(\alpha,\beta)}$. These are used in the sparsification process. We show how to sparsify path-clique graph classes when $S^{(\alpha,\beta)}$ is a path (Lemma 5.6) or tree (Lemma 5.7).

In Section 5.3 we show that path-clique hereditary graph classes have unbounded clique-width if the auxiliary star-graph $S^{(\alpha,\beta)}$ has an unbounded number of components or contains a rooted leafless tree (Theorem 5.13). We conjecture that a path-clique hereditary graph class has unbounded clique-width if and only if the auxiliary biclique-graph $B^{(\alpha,\beta)}$ contains an unbounded number of distinct cliques or bicliques (Conjecture 5.14).

In Section 5.4 we introduce the idea of words with *almost fixed support*, and show that such words exist (Lemma 5.15). In Theorem 5.19 we show that a hereditary class of graphs $\mathcal{H}^{(\alpha,\beta)}$ that has unbounded clique-width is minimal if α is almost periodic and has almost fixed support and the auxiliary star-graph $S^{(\alpha,\beta)}$ has bounded linear clique-width.

In Theorem 5.21 we prove that this alternative framework, like the framework in Chapter 2, contains uncountably many minimal classes.

1.12 Antichains of Unbounded Clique-width

As explained in Section 1.7, hereditary classes of unbounded clique-width that do not contain a minimal class must contain an antichain of unbounded clique-width, that is, a sequence of graphs $A_1, A_2, \ldots$ such that there is no pair of indices i, j such that $A_i \leqslant A_j$ or $A_j \leqslant A_i$ and $cw(A_i) \to \infty$ as $i \to \infty$.

In Section 1.10 we observed that certain sparse hereditary classes of unbounded tree-width (and therefore clique-width) do not contain a minimal class but instead contain a t-*basic obstruction*: that is, for arbitrarily large t, a subdivision of the $(t \times t)$-wall and the line graph of a subdivision of the $(t \times t)$-wall. Here, we exclude the hereditary classes, cliques and bicliques, since they each have bounded clique-width so are not t-basic obstructions to bounded clique-width.

We use the term 't-basic obstruction' to mean some graph structure that determines unboundedness of clique-width in certain hereditary graph classes that do not contain a minimal class or the other t-basic obstructions. In the case of t-sails we showed that nested word path-star graph classes did not contain a minimal class, were KKW-free, and had unbounded clique-width if and only if they contained arbitrarily large t-sails.

These t-basic obstructions must contain within them an antichain of unbounded clique-width, although these antichains may not be unique. Here we show an example of a $t \times t$-wall antichain in Figure 1.6 and an example of a t-sail antichain in Figure 1.7.

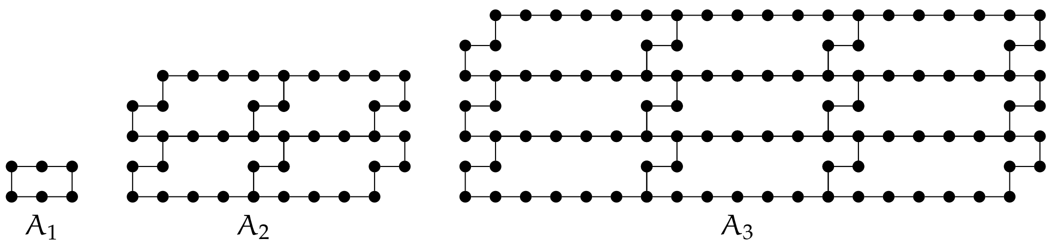


Figure 1.6: First three graphs in a $t \times t$ wall antichain of unbounded clique-width

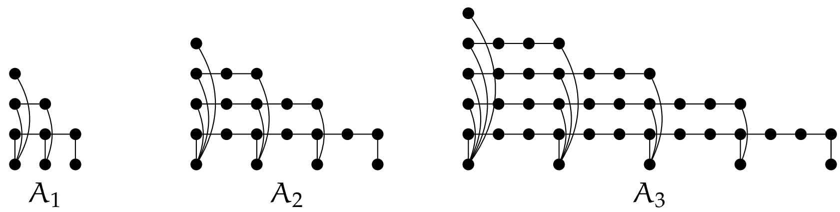


Figure 1.7: First three graphs in a t-sail antichain of unbounded clique-width

In Chapter 5, we introduce another t-basic obstruction – the clique t-sails (Definition 5.10). Although we do not formally prove it is a t-basic obstruction, it has the same necessary characteristics as a t-sail. Another way to generate further t-basic obstructions is to use Lemma 1.24, since we know that taking complements or bipartite complements of an antichain of unbounded clique-width will yield another antichain of unbounded clique-width. In Chapter 6 we introduce a further t-basic obstruction, the t-*clipper* (see Definition 6.2). An example is shown in Figure 1.8.

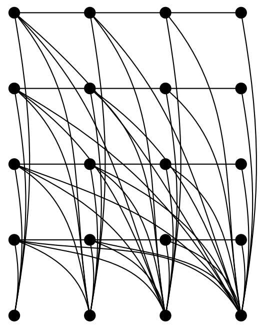


Figure 1.8: A 4-clipper

Korpelainen [42] defined a hereditary class of *permutation-partition graphs*, which we will refer to as *path-half graphs* for consistency, as the induced subgraphs of an infinite half graph, where one part of the bipartition is replaced by a linear forest (see full Definition 6.3). We revisit the class and prove that graphs in the class with large clique-width must contain a large t-clipper as an induced subgraph (Theorem 6.5). Furthermore, the class does not contain a minimal class (Theorem 6.8). Hence, the t-clipper is another t-basic obstruction.

Having given an example of an antichain of unbounded clique-width for t-clippers (see Figure 6.3) we suggest other similar objects that will also be t-basic obstructions to bounded clique-width: a clique t-clipper, a two-sided t-clipper and clique two-sided t-clipper (see Figure 6.4).

Chapter 2

Dense Classes : The Framework I - Unbounded Clique-width

2.1 Introduction

We create a framework for hereditary graph classes $\mathcal{G}^\delta$ built on a two-dimensional grid of vertices and edge sets defined by a triple $\delta = \{\alpha, \beta, \gamma\}$ of objects that define edges between consecutive columns, edges between non-consecutive columns (called bonds), and edges within columns. This framework captures a large family of minimal hereditary classes of graphs of unbounded clique-width, some previously identified and many new ones.

We show that a graph class $\mathcal{G}^\delta$ has unbounded clique-width if and only if a certain parameter $\mathcal{N}^\delta$ is unbounded. The parameter $\mathcal{N}^\delta$ is a property of triple $\delta = (\alpha, \beta, \gamma)$, and is a measure of the number of distinct neighbourhoods in a certain auxiliary graph.

We introduce new methods to proving that a hereditary class has unbounded clique-width, including the use of Ramsey theory.

2.2 Constructing the framework of $\mathcal{G}^\delta$ hereditary graph classes

The graph classes we consider are all formed by taking the set of finite induced subgraphs of an infinite graph defined on a grid of vertices. We start by defining an infinite empty graph $\mathcal{P}$, using Cartesian coordinates, with vertices

$$V(\mathcal{P}) = \{v_{i,j} : i, j \in \mathbb{N}\}.$$

We think of $\mathcal{P}$ as an infinite two-dimensional array in which $v_{i,j}$ represents the vertex in the i-th column (counting from the left) and the j-th row (counting from the bottom). Hence vertex $v_{1,1}$ is in the bottom left corner of the grid and the grid extends infinitely upwards and to the right. The i-th column of $\mathcal{P}$ is the set $C_i = \{v_{i,j} : j \in \mathbb{N}\}$, and the j-th row of $\mathcal{P}$ is the set $R_j = \{v_{i,j} : i \in \mathbb{N}\}$. Likewise, the collection of vertices in columns i to j is denoted $C_{[i,j]}$.

We will add edges to $\mathcal{P}$ using a triple δ of objects that define the edges between consecutive columns, edges between non-consecutive columns and edges within each column.

A *bond-set* β is a symmetric subset of $\{(x, y) \in \mathbb{N}^2, |x - y| > 1\}$. For a set $Q \subseteq \mathbb{N}$ we write β_Q to mean the subset of β-bonds $\{(x, y) \in \beta : x, y \in Q\}$. For instance, $\beta_{[i,j]} = \{(x, y) \in \beta : i \leqslant x, y \leqslant j\}$.

Let α be an infinite word over the alphabet $\{0, 1, 2, 3\}$, β be a bond set and γ be an infinite binary word. We refer to the three objects combined as a δ-*triple*, denoted $\delta = (\alpha, \beta, \gamma)$.

We define an infinite graph $\mathcal{P}^\delta$ with vertices $V(\mathcal{P})$ and with edges defined by δ as follows:

(a) α-*edges* between consecutive columns determined by the letters of the word α. For each $i = 1, 2, \ldots$, the edges between C_i and C_{i+1} are:

  (i) $\{(v_{i,j}, v_{i+1,j}) : j \in \mathbb{N}\}$ if $\alpha_i = 0$ (i.e. a matching);

  (ii) $\{(v_{i,j}, v_{i+1,k}) : j \neq k; j, k \in \mathbb{N}\}$ if $\alpha_i = 1$ (i.e. the bipartite complement of a matching);

  (iii) $\{(v_{i,j}, v_{i+1,k}) : j \geqslant k; j, k \in \mathbb{N}\}$ if $\alpha_i = 2$ (i.e. a half graph – see Section 1.2.1);

  (iv) $\{(v_{i,j}, v_{i+1,k}) : j < k; j, k \in \mathbb{N}\}$ if $\alpha_i = 3$ (i.e. the bipartite complement of a half graph).

(b) β-*edges* defined by the bond-set β such that $v_{i,x} \sim v_{j,y}$ for all $x, y \in \mathbb{N}$ when $(i, j) \in \beta$ (i.e. a complete bipartite graph between C_i and C_j), and

(c) γ-*edges* defined by the letters of the binary word γ such that for any $j, k \in \mathbb{N}$ we have $v_{i,j} \sim v_{i,k}$ if and only if $\gamma_i = 1$ (i.e. C_i forms a clique if $\gamma_i = 1$ and an independent set if $\gamma_i = 0$).

The hereditary graph class $\mathcal{G}^\delta$ is the set of all finite induced subgraphs of $\mathcal{P}^\delta$.

Any graph $G \in \mathcal{G}^\delta$ can be witnessed by an embedding $\phi(G)$ into the infinite graph $\mathcal{P}^\delta$. To simplify the presentation we will associate G with a particular embedding in $\mathcal{P}^\delta$ depending on the context. We will be especially interested in the induced subgraphs of

G that occur in consecutive columns: in particular, an α_j-*link* is the induced subgraph of G on the vertices of $G \cap C_{[j,j+1]}$, and will be denoted by $G_{[j,j+1]}$. More generally, an induced subgraph of G on the vertices of $G \cap C_{[j,k]}$ will be denoted $G_{[j,k]}$.

For $k \geqslant 2$ and $j \in \mathbb{N}$ we denote a triple $\delta_{[j,j+k-1]} = (\alpha_{[j,j+k-2]}; \beta_{[j,j+k-1]}; \gamma_{[j,j+k-1]})$ as a k-*factor* of δ. Thus for a graph $G \in \mathcal{G}^\delta$ with a particular embedding in $\mathcal{P}^\delta$, the induced subgraph $G_{[j,j+k-1]}$ has edges defined by the k-factor $\delta_{[j,j+k-1]}$.

We say that two k-factors $\delta_{[x,x+k]}$ and $\delta_{[y,y+k]}$ are the same if

(i) for all $i \in [0, k-1]$, $\alpha_{x+i} = \alpha_{y+i}$, and

(ii) for all $i, j \in [0, k]$, $(x+i, x+j) \in \beta$ if and only if $(y+i, y+j) \in \beta$, and

(iii) for all $i \in [0, k]$, $\gamma_{x+i} = \gamma_{y+i}$.

We say that a δ-triple is *recurrent* if every k-factor for $k \in \mathbb{N}$ occurs in it infinitely many times. We say that δ is *almost periodic* if for each k-factor $\delta_{[j,j+k-1]}$ of δ for $j, k \in \mathbb{N}$ there exists a constant $\mathcal{L}(\delta_{[j,j+k-1]})$ such that every factor of δ of length $\mathcal{L}(\delta_{[j,j+k-1]})$ contains $\delta_{[j,j+k-1]}$ as a factor.

2.3 $\mathcal{G}^\delta$ graph classes with unbounded clique-width

This section identifies which hereditary classes $\mathcal{G}^\delta$ have unbounded clique-width. We prove this is determined by a neighbourhood parameter $\mathcal{N}^\delta$ derived from a graph induced on any two rows of the graph $\mathcal{P}^\delta$. We show that $\mathcal{G}^\delta$ has unbounded clique-width if and only if $\mathcal{N}^\delta$ is unbounded (Theorem 2.17).

2.3.1 The two-row graph and $\mathcal{N}^\delta$

Informally, we show that boundedness of clique-width for a graph class $\mathcal{G}^\delta$ is determined by boundedness of the number of different neigbourhoods between the vertices in one row and the vertices of another row of $\mathcal{P}^\delta$ (it could be any two rows). To explain this precisely we need some further definitions, as follows:

Given a graph $G = (V, E)$ and a subset of vertices $U \subseteq V$, two vertices of U will be called $V \setminus U$-*similar* if they have the same neighbourhood in $V \setminus U$. Thus $V \setminus U$-similarity is an equivalence relation. The number of such equivalence classes of U in G will be denoted $\mu(G, U)$. A special case is when all the equivalence classes are singletons when we call U a *distinguished vertex set*.

A *distinguished pairing* $\{U, W\}$ of size r of a graph $G = (V, E)$ is a pair of vertex subsets $U = \{u_i : i \in [1, r]\} \subseteq V$ and $W = \{w_i : i \in [1, r]\} \subseteq V \setminus U$ with $|U| = |W| = r$ such that the vertices in U have pairwise different neighbourhoods in W (but not necessarily vice-versa). A distinguished pairing is *matched* if the vertices of U and W can be paired (u_i, w_i) so that $u_i \sim w_i$ for each i, and is *unmatched* if the vertices of U and W can be paired (u_i, w_i) so that $u_i \not\sim w_i$ for each i. Clearly the set U of a distinguished pairing $\{U, W\}$ is a distinguished vertex set of $G[U \cup W]$ which gives us the following:

Proposition 2.1. *If $\{U, W\}$ is a distinguished pairing of size r in graph G then $\mu(G[U \cup W], U) = r$.*

A *two-row graph* $T^\delta(Q) = (V, E)$ is the subgraph of $\mathcal{P}^\delta$ induced on the vertices $V = R_1(Q) \cup R_2(Q)$ where $R_1(Q) = \{v_{i,1} : i \in Q\}$ and $R_2(Q) = \{v_{j,2} : j \in Q\}$ for finite subset $Q \subseteq \mathbb{N}$.

We define the parameter $\mathcal{N}^\delta(Q) = \mu(T^\delta(Q), R_1(Q))$. See the examples in Figure 2.1.

Figure 2.1: $T^\delta([1, 8])$ where (on the left) $\delta = (0^\infty, \emptyset, 0^\infty)$ so $\mathcal{N}^\delta([1, 8]) = 1$ and (on the right) $\delta = (2^\infty, \emptyset, 0^\infty)$ so $\mathcal{N}^\delta([1, 8]) = 8$

Lemma 2.2. *For any fixed $j \in \mathbb{N}$, $\mathcal{N}^\delta([1, n])$ is bounded as $n \to \infty$ if and only if $\mathcal{N}^\delta([j, n])$ is bounded as $n \to \infty$.*

Proof. It is easy to see that if there exists N such that $\mathcal{N}^\delta([1, n]) < N$ for all $n \in \mathbb{N}$ then $\mathcal{N}^\delta([j, n]) < N$ for all $n \in \mathbb{N}$.

On the other hand, if $\mathcal{N}^\delta([j, n]) < N$ then $\mathcal{N}^\delta([j-1, n]) < 2N + 1$ since by adding the extra column each 'old' equivalence class could at most be split in two and there is one new vertex in each row. By induction we have $\mathcal{N}^\delta([1, n]) < 2^j N + \sum_{i=0}^{j-1} 2^i$ for all $n \in \mathbb{N}$. □

We will say $\mathcal{N}^\delta$ *is unbounded* if $\mathcal{N}^\delta([j, n])$ is unbounded as $n \to \infty$ for some fixed $j \in \mathbb{N}$. In many cases it is simple to check that $\mathcal{N}^\delta$ is unbounded – e.g. the following δ-triples have unbounded $\mathcal{N}^\delta$:

$$(1^\infty, \emptyset, 0^\infty), (2^\infty, \emptyset, 0^\infty), (3^\infty, \emptyset, 0^\infty), (0^\infty, \emptyset, 1^\infty)$$

In Lemma 2.14 we show that $\mathcal{N}^\delta$ is unbounded whenever α contains an infinite number of 2s or 3s.

2.3.2 Clique-width expression and colour partition for an $n \times n$ square graph

We will denote $H^{\delta}_{i,j}(m, n)$ as the $m(cols) \times n(rows)$ induced subgraph of $\mathcal{P}^{\delta}$ formed from the rectangular grid of vertices $\{v_{x,y} : x \in [i, i+m-1], y \in [j, j+n-1]\}$. See example in Figure 2.2.

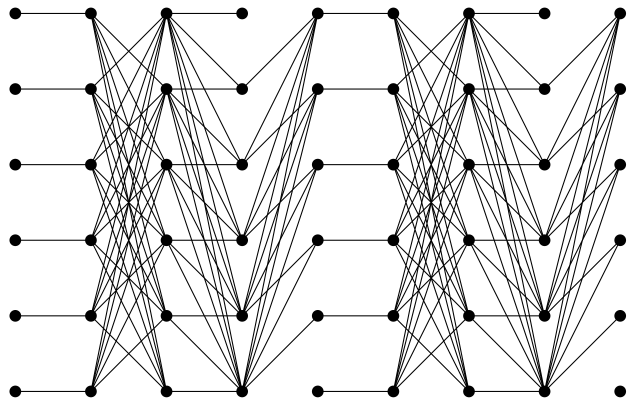


Figure 2.2: The graph $H^{\delta}_{1,1}(9,6)$ from the framework, where $\delta = (\alpha, \beta, \gamma)$ with $\alpha = 01230123\cdots$, $\beta = \emptyset$ and $\gamma = 0^{\infty}$

We can calculate a lower bound for the clique-width of the $n \times n$ square graph $H^{\delta}_{j,1}(n, n)$ (shortened to $H(n, n)$ when δ, j and 1 are clearly implied), by demonstrating a minimum number of labels needed to construct it using the allowed four graph operations, as follows.

Let τ be a clique-width expression defining $H(n, n)$ and the rooted tree representing τ. The subtree of tree(τ) rooted at a node $\oplus$ corresponds to a subgraph of $H(n, n)$. We can give this node a label, say a, so that $\oplus_a$ is the root and H_a the corresponding subgraph of $H(n, n)$.

We denote by $\oplus_{red}$ and $\oplus_{blue}$ the two children of $\oplus_a$ in tree(τ). Let us colour the vertices of H_{red} and H_{blue} red and blue, respectively, and all the other vertices in $H(n, n)$ white. Let colour(v) denote the colour of a vertex $v \in H(n, n)$ as described above, and label(v) denote the label of vertex v (if any) having applied the operations in tree(τ) up to node $\oplus_a$. (If v is white it is a vertex of $H(n, n)$ not in subgraph H_a and therefore it has either been created in a branch of tree(τ) not yet connected to node $\oplus_a$, or has not yet been created, in which case we say label(v) $= \epsilon$).

Our identification of a minimum number of labels needed to construct $H(n, n)$ relies on the following observation regarding this vertex colour partition.

Observation 2.3. *Suppose u_1, u_2, w are three vertices in $H(n, n)$ such that u_1 and u_2 are non-white, $u_1 \sim w$ but $u_2 \not\sim w$, and colour(w) $\neq$ colour(u_1). Then u_1 and u_2 must have different labels at node $\oplus_a$.*

This is true because the edge u_1w still needs to be created, whilst respecting the non-adjacency of u_2 and w. We now focus on sets of blue and sets of nonblue vertices (Equally, we could have chosen red-nonred). Observation 2.3 leads to the following key lemma which is the basis of much which follows.

Lemma 2.4. *For graph $H(n, n)$ let U and W be two disjoint vertex sets with induced subgraph $H = H(n, n)[U \cup W]$ such that $\mu(H, U) = r$. Then if the vertex colouring described above gives $\text{colour}(u) =$ blue for all $u \in U$ and $\text{colour}(w) \neq$ blue for all $w \in W$ then the clique-width expression τ requires at least r labels at node $\oplus_a$.*

Proof. Choose one representative vertex from each equivalence class in U. For any two such representatives u_1 and u_2 there must exist a w in W such that $u_1 \sim w$ but $u_2 \not\sim w$ (or vice versa). By Observation 2.3 u_1 and u_2 must have different labels at node $\oplus_a$. This applies to any pair of representatives u_1, u_2 and hence all r such vertices must have distinct labels. □

Note that from Proposition 2.1 a distinguished pairing gives us the sets U and W required for Lemma 2.4. The following lemmas identify structures in $H(n, n)$ that give us these distinguished pairings.

We denote by $H_{[y,y+1]}$ the α_y-link $H(n, n) \cap C_{[y,y+1]}$ where $y \in [j, j+n-2]$. We will refer to a (adjacent or non-adjacent) *blue-nonblue pair* to mean two vertices, one of which is coloured blue and one non-blue, such that they are in consecutive columns, where the blue vertex could be to the left or the right of the nonblue vertex. If we have a set of such pairs with the blue vertex on the same side (i.e. on the left or right) then we say the pairs in the set have the same *polarity*.

Lemma 2.5. *Suppose that $H_{[y,y+1]}$ contains a horizontal pair (b_1, b_2) of blue vertices and at least one nonblue vertex n_1, n_2 in each column, but not on the top or bottom row (see Figure 2.3).*

(a) If $\alpha_y \in \{0, 2, 3\}$ then $H_{[y,y+1]}$ contains a non-adjacent blue-nonblue pair.

(b) If $\alpha_y \in \{1, 2, 3\}$ then $H_{[y,y+1]}$ contains an adjacent blue-nonblue pair.

Proof. If $\alpha_y = 0$ then both (b_1, n_1) and (b_2, n_2) form a non-adjacent blue-nonblue pair (Figure 2.3 A). If $\alpha_y = 1$ then both (b_1, n_1) and (b_2, n_2) form an adjacent blue-nonblue pair (Figure 2.3 B).

If $\alpha_y \in \{2, 3\}$ and the nonblue vertices n_1 and n_2 in each column are either both above or both below the horizontal blue pair (b_1, b_2) then it can be seen that one of the pairs

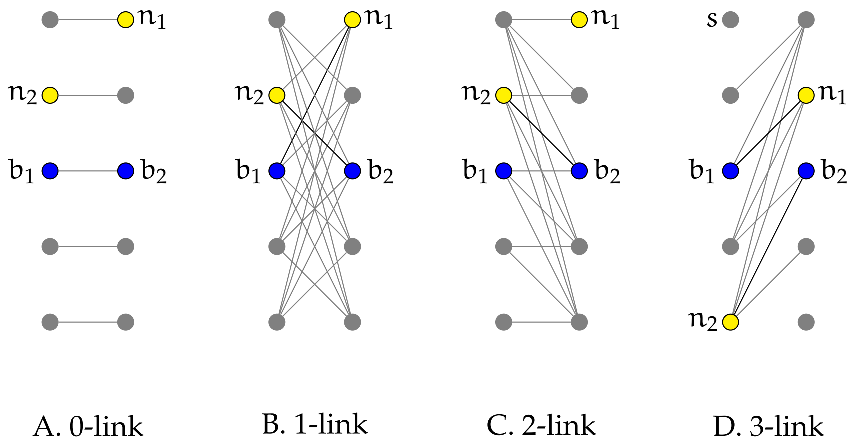


Figure 2.3: Horizontal blue-blue pair in $H_{[y,y+1]}$ (nonblue vertices in yellow)

(b_1, n_1) or (b_2, n_2) forms an adjacent blue-nonblue pair and the other forms a non-adjacent blue-nonblue pair (Figure 2.3 C). If the nonblue vertices in each column are either side of the blue pair (one above and one below) then the pairs (b_1, n_1) and (b_2, n_2) will both be adjacent (or non-adjacent) blue-nonblue pairs (See Figure 2.3 D). In this case we need to appeal to a 5-th vertex s which will form a non-adjacent (or adjacent) set with either n_1 or b_2 depending on its colour. Thus we always have both a non-adjacent and adjacent blue-non-blue pair when $\alpha_y \in \{2, 3\}$. □

Lemma 2.6. *Suppose* $H_{[y,y+1]}$ *contains a horizontal blue-nonblue pair of vertices* (b_1, n_1)*, not the top or bottom row, and at least one nonblue vertex* n_2 *in the same column as* b_1*. Then* $H_{[y,y+1]}$ *contains both an adjacent and a non-adjacent blue-nonblue pair of vertices, irrespective of the value of* α_y *(see Figure 2.4).*

Proof. If $\alpha_y \in \{0, 2\}$ then the horizontal blue-nonblue pair (b_1, n_1) is adjacent, and given a nonblue vertex n_2 in the same column as b_1, we can find a vertex s in the same column as n_1 that forms a non-adjacent pairing with either b_1 or n_2 depending on its colour (See Figure 2.4 A and C). If $\alpha_y \in \{1, 3\}$ then the horizontal blue-nonblue pair (b_1, n_1) is non-adjacent, and given a nonblue vertex n_2 in the same column as b_1, we can find a vertex s in the same column as n_1 that forms an adjacent pairing with either b_1 or n_2 depending on its colour (See Figure 2.4 B and D). □

Lemma 2.7. *Suppose* $H_{[y,y+1]}$ *contains* $r \geqslant 3$ *horizontal blue-nonblue pairs of vertices with the same polarity,* $(b_1, n_1), \ldots, (b_r, n_r)$ *(see Figure 2.5). Then, irrespective of the value of* α_y*, it contains*

(a) *a matched distinguished pairing* $\{U, W\}$ *of size* $r - 1$ *such that* $\text{colour}(u) =$ *blue for all* $u \in U$ *and* $\text{colour}(w) \neq$ *blue for all* $w \in W$*, and*

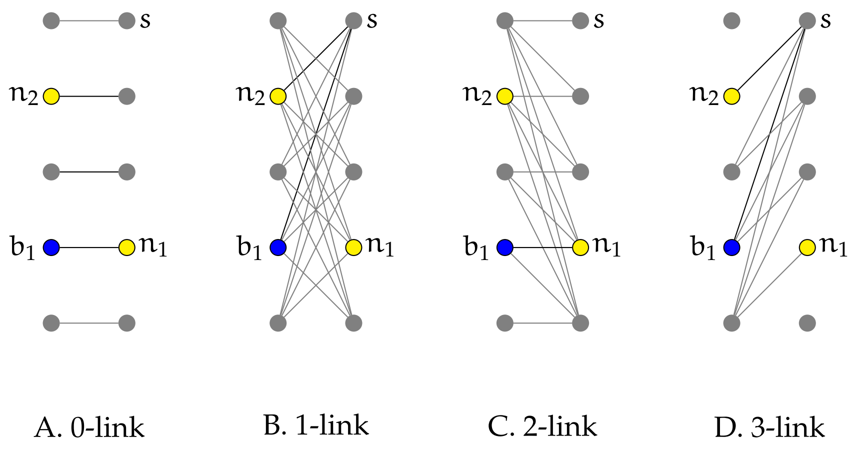


Figure 2.4: Horizontal blue-nonblue pair in $H_{[y,y+1]}$ (nonblue vertices in yellow)

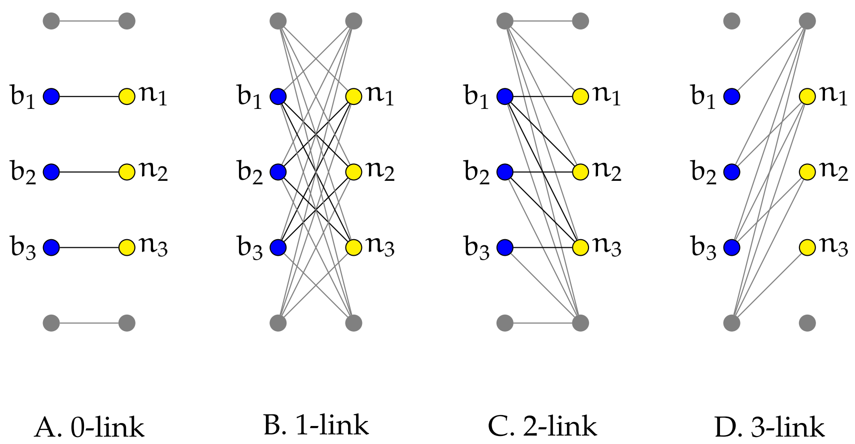


Figure 2.5: 3 horizontal blue-nonblue pairs in $H_{[y,y+1]}$ (nonblue vertices in yellow)

(b) an unmatched distinguished pairing $\{U', W'\}$ *of size* $r-1$ *such that* $\text{colour}(u') =$ *blue for all* $u' \in U'$ *and* $\text{colour}(w') \neq$ *blue for all* $w' \in W'$.

Proof. This is easily observable from Figure 2.5 for $r = 3$. If we set $U = \{b_1, b_2\}$, $W = \{n_1, n_2\}$, $U' = \{b_2, b_3\}$ and $W' = \{n_1, n_2\}$ then one of $\{U, W\}$ and $\{U', W'\}$ is a matched distinguished pairing of size 2 and the other is an unmatched distinguished pairing of size 2, irrespective of the value of α_y. Simple induction establishes this for all $r \geqslant 3$. □

A *couple set* P is a subset of $\mathbb{N}$ such that if $x, y \in P$ then $|x - y| > 2$. Such a set is used to identify sets of links that have no α-edges between them. We say that a pair (x, y) of elements of P is β-*dense* if both $(x, y+1)$ and $(x+1, y)$ are in β and they are β-*sparse* when neither of these bonds is in β.

We say the bond-set β is *sparse* in P if every pair from P is β-sparse and is *not sparse* in P if there are no β-sparse pairs in P. Likewise, β is *dense* in P if every pair from P

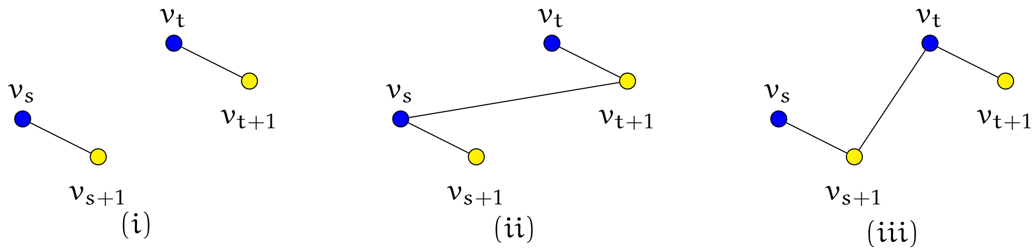


Figure 2.6: Adjacent blue-nonblue vertex pairs, β not dense (nonblue vertices in yellow)

is β-dense and is *not dense* in P if there are no β-dense pairs in P. Clearly it is possible for two elements from P to be neither β-sparse nor β-dense (i.e. when only one of the required bonds is in β). These ideas are used to identify matched and unmatched distinguished pairings (see Lemmas 2.8 and 2.9).

In Lemmas 2.5, 2.6 and 2.7 we identified blue-nonblue pairs within a particular link $H_{[y,y+1]}$. The next two lemmas identify distinguished pairings across link-sets. Let $P \subset [j, j+n-2]$ be a couple set of size r with corresponding α_y-links $H_{[y,y+1]} \leqslant H(n,n)$ for each $y \in P$.

Lemma 2.8. *If β is not dense in P and each $H_{[y,y+1]}$ for $y \in P$ has an adjacent blue-nonblue pair with the same polarity, then we can combine these pairs to form a matched distinguished pairing $\{U, W\}$ of size r where the vertices of U are blue and the vertices of W nonblue.*

Proof. Suppose $s, t \in P$ such that (v_s, v_{s+1}) and (v_t, v_{t+1}) are two adjacent blue-nonblue pairs in different links, with $v_s, v_t \in U$ and $v_{s+1}, v_{t+1} \in W$. Consider the two possible β bonds (v_s, v_{t+1}) and (v_{s+1}, v_t). If neither of these bonds exist then v_s is distinguished from v_t by both v_{s+1} and v_{t+1} (see Figure 2.6 (i)). If one of these bonds exists then v_s is distinguished from v_t by either v_{s+1} or v_{t+1} (see Figure 2.6 (ii) and (iii)). Both bonds cannot exist as β is not dense in P. Note that the bonds (v_s, v_t) and (v_{s+1}, v_{t+1}) are not relevant in distinguishing v_s from v_t since, if they exist, they connect blue to blue and nonblue to nonblue.

So any two blue vertices $v_s, v_t \in U$ are distinguished by the two nonblue vertices $v_{s+1}, v_{t+1} \in W$ and hence $\{U, W\}$ is a matched distinguished pairing of size r. □

Lemma 2.9. *If β is not sparse in P and each $H_{[y,y+1]}$ has a non-adjacent blue-nonblue pair with the same polarity, then we can combine these pairs to form an unmatched distinguished pairing $\{U, W\}$ of size r where the vertices of U are blue and the vertices of W nonblue.*

Proof. This is very similar to the proof of Lemma 2.8 and is demonstrated in Figure 2.7. □

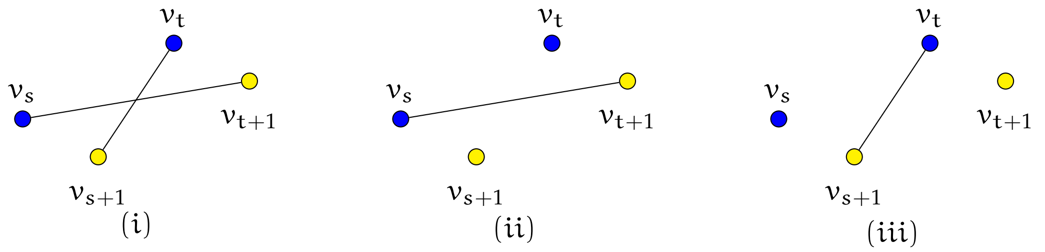


Figure 2.7: Non-adjacent blue-nonblue vertex pairs, β not sparse (nonblue vertices in yellow)

2.3.3 Two colour partition cases to consider

Having identified structures that give us a lower bound on labels required for a clique-width expression for $H(n, n)$, we now apply this knowledge to the following subtree of tree(τ).

Let $\oplus_a$ be the lowest node in tree(τ) (that is, furthest from the root) such that H_a contains all the vertices in rows 2 to $(n-1)$ in some column of $H(n, n)$. We reserve rows 1 and n so that we may apply Lemmas 2.5 and 2.6.

Thus $H(n, n)$ contains at least one column where vertices in rows 2 to $(n-1)$ are non-white but no column has entirely blue or red vertices in rows 2 to $(n-1)$ because otherwise $\oplus_a$ would not be the lowest node in tree(τ) such that H_a contains all the vertices in rows 2 to $(n-1)$ in some column of $H(n, n)$. Let C_b be a non-white column. Without loss of generality we can assume that the number of blue vertices in column C_b between rows 2 and $(n-1)$ is at least $(n/2) - 1$ otherwise we could swap red for blue.

Now consider rows 2 to $(n-1)$. We have two possible cases:

Case 1 Either none of the rows with a blue vertex in column C_b has blue vertices in every column to the right of C_b, or none of the rows with a blue vertex in column C_b has blue vertices in every column to the left of C_b. Hence, we have at least $\lceil n/2 \rceil - 1$ rows that have a horizontal blue-nonblue pair with the same polarity.

Case 2 One row R_r has a blue vertex in column C_b and blue vertices in every column to the right of C_b and one row R_l has a blue vertex in column C_b and blue vertices in every column to the left of C_b. Hence, either on row R_r or row R_l, we must have a horizontal set of consecutive blue vertices of size at least $\lceil n/2 \rceil + 1$.

To prove unboundedness of clique-width we will show that for any $r \in \mathbb{N}$ we can find an $n \in \mathbb{N}$ so that any clique-width expression τ for $H(n, n)$ requires at least r labels in tree(τ), whether this is a 'Case 1' or 'Case 2' scenario.

To address both cases we will need the following classic result:

Theorem 2.10 (Ramsey [56] and see Diestel [30])**.** *For every* $r \in \mathbb{N}$*, every graph of order at least* 2^{2r-3} *contains either* K_r *or* $\overline{K_r}$ *as an induced subgraph.*

We handle first Case 1, for all values of $\delta = (\alpha, \beta, \gamma)$.

Lemma 2.11. *For any* $\delta = (\alpha, \beta, \gamma)$ *and any* $r \in \mathbb{N}$*, if* $n \geqslant 9 \times 2^{4r-1}$ *and* τ *is a clique-width expression for* $H(n, n)$ *that results in Case 1 at node* $\oplus_a$*, then* τ *requires at least* r *labels to construct* $H(n, n)$*.*

Proof. In Case 1 we have, without loss of generality, at least $\lceil n/2 \rceil - 1$ horizontal blue-nonblue vertex pairs but we don't know which links these fall on.

If there are at least $\sqrt{n/2}$ such pairs on the same link then using Lemma 2.7 we have a matched distinguished pairing $\{U, W\}$ of size $\sqrt{n/2} - 1 > r$ such that $\text{colour}(u) = \text{blue}$ for all $u \in U$ and $\text{colour}(w) \neq \text{blue}$ for all $w \in W$.

If there is no link with $\sqrt{n/2}$ such pairs then at least $\sqrt{n/2}$ different links contain such a pair. From Lemma 2.6 each such link contains both an adjacent and non-adjacent blue-nonblue pair. It follows from the pigeonhole principle that there is a subset of these of size $\sqrt{n/2}/4$ where the adjacent blue-nonblue pairs have the same polarity and also the non-adjacent blue-nonblue pairs have the same polarity. We use this subset (Note, the following argument applies whether the blue vertex is on the left or right for the adjacent and non-adjacent pairs). If we take the index of the first column in each link in the mentioned subset, and then take every third one of these, we have a couple set P where $|P| \geqslant \sqrt{n/2}/12$, with corresponding link set $S_L = \{H_{[y,y+1]} : y \in P\}$, such that the adjacent blue-nonblue pair in each link has the same polarity and the non-adjacent blue-nonblue pair in each link has the same polarity.

Define the graph G_P so that $V(G_P) = P$ and for $x, y \in V(G_P)$ we have $x \sim y$ if and only if they are β-dense (see definition on page 34). From Theorem 2.10 for any r, as $|P| \geqslant \sqrt{n/2}/12 \geqslant 2^{2r-3}$ then there exists a couple set $Q \subseteq P$ such that G_Q is either K_r or $\overline{K_r}$.

If G_Q is $\overline{K_r}$, it follows that β is not dense in Q, and S_L contains a link set of size r corresponding to the couple set Q where each link has an adjacent blue-nonblue pair with the same polarity. Applying Lemma 2.8 this gives us a matched distinguished pairing $\{U, W\}$ of size r such that $\text{colour}(u) = \text{blue}$ for all $u \in U$ and $\text{colour}(w) \neq \text{blue}$ for all $w \in W$.

If G_Q is K_r, it follows that β is dense and hence not sparse in Q, and S_L contains a link set of size r corresponding to the couple set Q where each link has a non-adjacent blue-nonblue pair with the same polarity. Applying Lemma 2.9 this gives us an unmatched distinguished pairing $\{U, W\}$ of size r such that colour(u) = blue for all $u \in U$ and colour$(w) \neq$ blue for all $w \in W$.

Both cases give us a distinguished pairing $\{U, W\}$ of size r such that colour(u) = blue for all $u \in U$ and colour$(w) \neq$ blue for all $w \in W$. Hence, from Lemma 2.4 τ uses at least r labels to construct $H(n, n)$. □

2.3.4 When α has an infinite number of 2s or 3s

For Case 2 we must consider different values for α separately. We denote $m_{23}(n)$ to be the total number of 2s and 3s in $\alpha_{[1,n-1]}$.

Lemma 2.12. *For any triple $\delta = (\alpha, \beta, \gamma)$ and any $r \in \mathbb{N}$, if $m_{23}(n) \geqslant 3 \times 2^{2r}$ and τ is a clique-width expression for $H(n, n)$ that results in Case 2 at node $\oplus_a$, then τ requires at least r labels to construct $H(n, n)$.*

Proof. Remembering that C_b is the non-white column, without loss of generality we can assume that there are at least $(m_{23}(n)/2)$ 2- or 3-links to the right of C_b, since otherwise we could reverse the order of the columns. In Case 2 each link has a horizontal blue-blue vertex pair with at least one nonblue vertex in each column, so using Lemma 2.5 we have both an adjacent and non-adjacent blue-nonblue pair in each of these links.

It follows from the pigeonhole principle that there is a subset of these of size $(m_{23}(n)/8)$ where the adjacent blue-nonblue pairs have the same polarity and also the non-adjacent blue-nonblue pairs have the same polarity. We use this subset. If we take the index of the first column in each link in the mentioned subset, and then take every third one of these, we have a couple set P where $|P| \geqslant (m_{23}(n)/24)$, with corresponding link set $S_L = \{H_{[y,y+1]} : y \in P\}$, such that the adjacent blue-nonblue pair in each link has the same polarity and the non-adjacent blue-nonblue pair in each link has the same polarity.

As in the proof of Lemma 2.11, we define a graph G_P so that $V(G_P) = P$ and for $x, y \in V(G_P)$ we have $x \sim y$ if and only if they are β-dense. From Theorem 2.10 for any r, as $|P| \geqslant (m_{23}(n))/24) \geqslant 2^{2r-3}$ then there exists a couple set $Q \subseteq P$ such that G_Q is either K_r or $\overline{K_r}$.

We now proceed in an identical way to Lemma 2.11 to show that we can always construct a distinguished pairing $\{U, W\}$ of size r such that colour(u) = blue for all $u \in U$

and colour$(w) \neq$ blue for all $w \in W$. Hence, from Lemma 2.4 τ uses at least r labels to construct $H(n, n)$. □

Corollary 2.13. *For any triple $\delta = (\alpha, \beta, \gamma)$ such that α has an infinite number of 2s or 3s the hereditary graph class $\mathcal{G}^\delta$ has unbounded clique-width.*

Proof. This follows directly from Lemma 2.11 for Case 1 and Lemma 2.12 for Case 2, since for any $r \in \mathbb{N}$ we can choose n big enough so that $n \geqslant 9 \times 2^{4r-1}$ and $m_{23}(n) \geqslant 3 \times 2^{2r}$ so that whether we are in Case 1 or Case 2 at node $\oplus_a$ we require at least r labels for any clique-width expression for $H(n, n)$. □

We are aiming to state our result in terms of unbounded $\mathcal{N}^\delta$ so we also require the following.

Lemma 2.14. *For any triple $\delta = (\alpha, \beta, \gamma)$ such that α has an infinite number of 2s or 3s the parameter $\mathcal{N}^\delta$ is unbounded.*

Proof. If there is an infinite number of 2s in α we can create a couple set P of any required size such that $\alpha_x = 2$ for every $x \in P$, so that in the two-row graph (see Section 2.3.1) $v_{x,1} \not\sim v_{x+1,2}$ and $v_{x,2} \sim v_{x+1,1}$ (i.e. we have both an adjacent and non-adjacent pair in the α_x-link).

We now apply the same approach as in Lemmas 2.11 and 2.12, applying Ramsey theory to the graph G_P defined in the same way as before. Then for any r we can set $|P| \geqslant 2^{2r-3}$ so that there exists a couple set $Q \subseteq P$ where G_Q is either K_r or $\overline{K_r}$.

If G_Q is $\overline{K_r}$ it follows that β is not dense in Q. So for any $x, y \in Q$, $v_{x+1,1}$ and $v_{y+1,1}$ have different neighbourhoods in $R_2(Q)$ since they are distinguished by either $v_{x,2}$ or $v_{y,2}$. Hence, if n is the highest natural number in Q then $\mathcal{N}^\delta([1, n+1]) \geqslant r$.

If G_Q is K_r it follows that β is not sparse in Q. So for any $x, y \in Q$, $v_{x,1}$ and $v_{y,1}$ have different neighbourhoods in $R_2(Q)$ since they are distinguished by either $v_{x+1,2}$ or $v_{y+1,2}$. Hence, $\mathcal{N}^\delta([1, n+1]) \geqslant r$.

Either way, we have $\mathcal{N}^\delta([1, n+1]) \geqslant r$, but r can be arbitrarily large, so $\mathcal{N}^\delta$ is unbounded.

A similar argument applies if there is an infinite number of 3s. □

2.3.5 When α has a finite number of 2s and 3s

If α contains only a finite number of 2s and 3s then there exists $J \in \mathbb{N}$ such that $\alpha_j \in \{0,1\}$ for $j > J$. In Case 2, where we have a part-row of consecutive blue vertices, we are interested in the adjacencies of these blue vertices to the nonblue vertices in each column. Although the nonblue vertices could be in any row, in fact, if α is over the alphabet $\{0,1\}$, the row index of the nonblue vertices does not alter the blue-nonblue adjacencies.

In Case 2, let Q be the set of column indices of the horizontal set of consecutive blue vertices in row R_r of $H(n,n)$ and let $U_1 = \{v_{i,r} : i \in Q\}$ be this horizontal set of blue vertices. Let $U_2 = \{u_j : j \in Q\}$ be the corresponding set of nonblue vertices such that $u_j \in C_j$. We have the following:

Lemma 2.15. *In Case 2, with U_1 and U_2 defined as above, if α is a word over the alphabet $\{0,1\}$ then for any $i,j \in Q$, $v_{i,r} \sim u_j$ in $\mathcal{P}^\delta$ if and only if $v_{i,1} \sim v_{j,2}$ in the two-row graph $T^\delta(Q)$.*

Proof. Considering the vertex sets $U_1 \cup U_2$ of $\mathcal{P}^\delta$ and $R_1(Q) \cup R_2(Q)$ of $T^\delta(Q)$ (see Section 2.3.1) we have:

(a) For $i = j$ both $v_{j,r} \sim u_j$ and $v_{j,1} \sim v_{j,2}$ if and only if $\gamma_j = 1$.

(b) For $|i-j| > 1$ both $v_{i,r} \sim u_j$ and $v_{i,1} \sim v_{j,2}$ if and only if $(i,j) \in \beta$.

(c) For $j = i+1$ both $v_{i,r} \sim u_j$ and $v_{i,1} \sim v_{j,2}$ if and only if $\alpha_i = 1$.

Hence $v_{i,r} \sim u_j$ if and only if $v_{i,1} \sim v_{j,2}$. □

Lemma 2.16. *If $\delta = (\alpha, \beta, \gamma)$ where α is an infinite word over the alphabet $\{0,1,2,3\}$ with a finite number of 2s and 3s, then the hereditary graph class $\mathcal{G}^\delta$ has unbounded clique-width if and only if $\mathcal{N}^\delta$ is unbounded.*

Proof. First, we prove that $\mathcal{G}^\delta$ has unbounded clique-width if $\mathcal{N}^\delta$ is unbounded.

As α has a finite number of 2s and 3s there exists a $J \in \mathbb{N}$ such that $\alpha_j \in \{0,1\}$ if $j > J$.

As $\mathcal{N}^\delta$ is unbounded this means that from Lemma 2.2 for any $r \in \mathbb{N}$ there exist $N_1, N_2 \in \mathbb{N}$ such that, setting $Q_1 = [J+1, J+N_1]$ and $Q_2 = [J+N_1+1, J+N_1+N_2]$, then $\mathcal{N}^\delta(Q_1) \geqslant r$ and $\mathcal{N}^\delta(Q_2) \geqslant r$.

Denote the $n \times n$ graph $H'(n,n) = H^\delta_{J+1,1}(n,n) \in \mathcal{G}^\delta$. As described in Section 2.3.3 we again consider the two possible cases for a clique-width expression τ for $H'(n,n)$ at a node $\oplus_a$ which is the lowest node in tree(τ) such that H_a contains a column of $H'(n,n)$.

Case 1 is already covered by Lemma 2.11 for $n \geqslant 9 \times 2^{4r-1}$.

In Case 2, one row R_r of $H'(N_1 + N_2, N_1 + N_2)$ has a blue vertex in column C_b and blue vertices in every column to the right of C_b and one row R_l has a blue vertex in column C_b and blue vertices in every column to the left of C_b.

If $b \leqslant J + N_1$ then consider the graph to the right of C_b. We know every column has a blue vertex in row R_r and a non-blue vertex in a row other than R_r. The column indices to the right of C_b includes Q_2. It follows from Lemma 2.15 that in the columns whose indices belong to Q_2 the neighbourhoods of the blue set (the mentioned blue vertices) to the non-blue set, are identical to the neighbourhoods in graph $T^\delta(Q_2)$ between the vertex sets $R_1(Q_2)$ and $R_2(Q_2)$.

On the other hand if $b > J + N_1$ we can make an identical claim for the graph to the left of C_b which now includes the column indices for Q_1. It follows from Lemma 2.15 that the neighbourhoods of the blue set to the non-blue set are identical to the neighbourhoods in graph $T^\delta(Q_1)$ between the vertex sets $R_1(Q_1)$ and $R_2(Q_1)$.

As both $\mathcal{N}^\delta(Q_1) = \mu(T^\delta(Q_1), R_1(Q_1)) \geqslant r$ and $\mathcal{N}^\delta(Q_2) = \mu(T^\delta(Q_2), R_2(Q_2)) \geqslant r$ it follows from Lemma 2.4 that any clique-width expression for $H'(n, n)$ with $n \geqslant (N_1 + N_2)$ resulting in Case 2 requires at least r labels.

For any $r \in \mathbb{N}$ we can choose n big enough so that $n \geqslant \max\{9 \times 2^{4r-1}, (N_1 + N_2)\}$ so that whether we are in Case 1 or Case 2 at node $\oplus_a$ we require at least r labels for any clique-width expression for $H'(n, n)$. Hence, $\mathcal{G}^\delta$ has unbounded clique-width if $\mathcal{N}^\delta$ is unbounded.

Secondly, suppose that $\mathcal{N}^\delta$ is bounded, so that there exists $N \in \mathbb{N}$ such that $\mathcal{N}^\delta([J + 1, n]) = \mu(T^\delta([J + 1, n]), R_1([J + 1, n])) < N$ for all $n > J$.

We claim $lcw(\mathcal{G}^\delta) \leqslant 2J + N + 2$. For we can create a linear clique-width expression using no more than $2J + N + 2$ labels that constructs any graph in $\mathcal{G}^\delta$ row by row, from bottom to top and from left to right.

For any graph $G \in \mathcal{G}^\delta$ let it have an embedding in the grid $\mathcal{P}$ between columns 1 and $M > J$.

We will use the following set of $2J + N + 2$ labels:

- 2 *current vertex labels*: a_1 and a_2;
- J *current row labels for first* J *columns*: $\{c_y : y = 1, \ldots, J\}$;
- J *previous row labels for first* J *columns*: $\{p_y : y = 1, \ldots, J\}$;

- N *partition labels*: $\{s_y : y = 1, \ldots, N\}$.

We allocate a default partition label s_y to each column of $G_{[J+1,M]}$ according to the $R_2([J+1, M])$-similar equivalence classes of the vertex set $R_1([J+1, M])$ in $T^{\delta}([J+1, M])$. There are at most N partition sets $\{S_y\}$ of $R_1([J + 1, M])$, and if vertex $v_{i,1}$ is in S_y, $1 \leqslant y \leqslant N$, then the default partition label for vertices in column i is s_y. It follows that for two default column labels, s_x and s_y, vertices in columns with label s_y are either all adjacent to vertices in columns with label s_x or they are all non-adjacent (except the special case of vertices in consecutive columns and the same row, which will be dealt with separately in our clique-width expression).

Carry out the following row-by-row linear iterative process to construct each row j, starting with row 1.

(i) Construct the first J vertices in row j, label them c_1 to c_J and build any edges between them as necessary.

(ii) Insert required edges from each vertex labelled $c_1, \ldots, c_J$ to vertices in lower rows in columns 1 to J. This is possible because the vertices in lower rows in column i ($1 \leqslant i \leqslant J$) all have label p_i and have the same adjacency with the vertices in the current row.

(iii) Relabel vertices labelled $c_1, \ldots, c_J$ to $p_1, \ldots, p_{J-1}, a_2$ respectively.

(iv) Construct and label subsequent vertices in row j (columns $J+1$ to M), as follows.

  (a) Construct the next vertex in column i and label it a_1 (or a_2).

  (b) If $\alpha_{i-1} = 0$ then insert an edge from the current vertex $v_{i,j}$ (label a_1) to the previous vertex $v_{i-1,j}$ (label a_2).

  (c) Insert edges to vertices that are adjacent as a result of the partition $\{S_y\}$ described above. This is possible because all previously constructed vertices with a particular default partition label s_y are either all adjacent or all non-adjacent to the current vertex.

  (d) Insert edges from the current vertex to vertices labelled p_j ($1 \leqslant j \leqslant J$) as necessary.

  (e) Relabel vertex $v_{i,j-1}$ to its default partition label s_y.

  (f) Create the next vertex in row i and label it a_2 (or a_1 alternating).

(v) When the end of the row is reached, repeat for the next row.

Hence we can construct any graph in the class with at most $2J + N + 2$ labels so the clique-width of $\mathcal{G}^\delta$ is bounded if $\mathcal{N}^\delta$ is bounded. □

Corollary 2.13, Lemma 2.14 and Lemma 2.16 give us the following:

Theorem 2.17. *For any triple $\delta = (\alpha, \beta, \gamma)$ the hereditary graph class $\mathcal{G}^\delta$ has unbounded clique-width if and only if $\mathcal{N}^\delta$ is unbounded.*

We carry forward this result to Chapter 3 by denoting Δ as the set of all δ-triples for which the class $\mathcal{G}^\delta$ has unbounded clique-width.

Chapter 3

Dense Classes : The Framework II - Minimal Classes

3.1 Introduction

We show that a hereditary graph class $\mathcal{G}^\delta$ with unbounded clique-width, taken from the framework of Chapter 2, is minimal of unbounded clique-width (and, indeed, minimal of unbounded linear clique-width) if another parameter $\mathcal{M}^\beta$ is bounded, and also δ has defined recurrence characteristics. Like the parameter $\mathcal{N}^\delta$, $\mathcal{M}^\beta$ is a property of triple $\delta = (\alpha, \beta, \gamma)$, and is a measure of the number of distinct neighbourhoods in a defined auxiliary graph.

Minimality is demonstrated by providing explicit bounded (linear) clique-width expressions for graphs in proper subclasses of minimal classes of unbounded clique-width. We introduce new methods to proving that a hereditary class is minimal, including a partition of the vertices referred to as 'veins' and 'slices'.

We demonstrate the framework contains an uncountable number of distinct minimal classes and that classes defined by recurrent but not almost periodic words exist.

3.2 $\mathcal{G}^\delta$ graph classes that are minimal of unbounded clique-width

To show that for some $\delta \in \Delta$ the class $\mathcal{G}^\delta$ is a minimal class of unbounded clique-width we must show that any proper hereditary subclass $\mathcal{C}$ has bounded clique-width. If $\mathcal{C}$ is a hereditary graph class such that $\mathcal{C} \subsetneq \mathcal{G}^\delta$ then there must exist a non-trivial finite forbidden graph F that is in $\mathcal{G}^\delta$ but not in $\mathcal{C}$. In turn, this graph F must be an induced subgraph of some $H^\delta_{j,1}(k,k)$ for some j and $k \in \mathbb{N}$, and thus $\mathcal{C} \subseteq \text{Free}(H^\delta_{j,1}(k,k))$.

Our strategy for proving that an arbitrary graph G in a proper hereditary subclass of $\mathcal{G}^\delta$ has bounded linear clique-width (and hence bounded clique-width) is to define an algorithm to create a linear clique-width expression that allows us to recycle labels so that we can put a bound on the total number of labels required, however many vertices there are in G.

For a minimal class, δ must be recurrent. For suppose it is not recurrent then it contains some k-factor $\delta_{[j,j+k-1]}$ that either does not repeat, or repeats only a finite number of times. There must be some finite $n \in \mathbb{N}$ such that there are no occurrences of the k-factor $\delta_{[j,j+k-1]}$ after the n-th letter in δ. Let $\epsilon = \delta_{[n,\infty)}$ so that $\mathcal{G}^\epsilon \subset \mathcal{G}^\delta$. Then $\mathcal{N}^\epsilon$ is unbounded, by Theorem 2.17 $\mathcal{G}^\epsilon$ has unbounded clique-width, and so $\mathcal{G}^\delta$ is not minimal, a contradiction. Therefore, we will only consider recurrent δ in this chapter.

3.2.1 The bond-graph

To study minimality we will use the following graph class. A *bond-graph* $B^\beta(Q) = (V,E)$ for finite $Q \subseteq \mathbb{N}$ has vertices $V = Q$ and edges $E = \beta_Q$.

Let $\mathcal{B}^\beta = \{B^\beta(Q) : Q \subseteq \mathbb{N} \text{ finite}\}$. Note that $\mathcal{B}^\beta$ is a hereditary subclass of $\mathcal{G}^\delta$ because

(a) if $Q' \subseteq Q$ then $B^\beta(Q')$ is also a bond-graph, and

(b) $B^\beta(Q)$ is an induced subgraph of $\mathcal{P}^\delta$ since if $Q = \{y_1, y_2, \ldots, y_n\}$ with $y_1 < y_2 < \cdots < y_n$ then it can be constructed from $\mathcal{P}^\delta$ by taking one vertex from each column y_j in turn such that there is no α or γ edge to previously picked vertices.

We define a parameter (for $n \geqslant 2$)

$$\mathcal{M}^\beta(n) = \sup_{m<n} \mu(B^\beta([1,n]),[1,m]).$$

The bond-graphs can be characterised as the sub-class of graphs on a single row (although missing the α-edges) with the parameter $\mathcal{M}^\beta$ measuring the number of distinct neighbourhoods between intervals of a single row (See the examples in Figure 3.1).

We will say that the bond-set β has *bounded* $\mathcal{M}^\beta$ if there exists M such that $\mathcal{M}^\beta(n) < M$ for all $n \in \mathbb{N}$.

The following proposition will prove useful later in creating linear clique-width expressions.

Figure 3.1: Examples of bond-graphs, $B^\beta([1,8])$, where (on the left) $\mathcal{M}^\beta(8) = 3$ and (on the right) $\mathcal{M}^\beta(8) = 4$

Proposition 3.1. *Let $n, m, m' \in \mathbb{N}$ satisfy $m < m' < n$. Then for graph $B^\beta([1,n])$, in any partition of $[1,m]$ into $[m+1,n]$-similar sets $\{S_i : 1 \leqslant i \leqslant k\}$ and $[1,m']$ into $[m'+1,n]$-similar sets $\{S'_j : 1 \leqslant j \leqslant k'\}$ for every $\ell \in [1,k]$ there exists $\ell' \in [1,k']$ such that $S_\ell \subseteq S'_{\ell'}$.*

Proof. As two vertices x and y in S_ℓ have the same neighbourhood in $[m+1,n]$ it follows they have the same neighbourhood in $[m'+1,n]$ since $m < m'$ so x and y must sit in the same $[m'+1,n]$-similar set $S'_{\ell'}$ for some $\ell' \in [1,k']$. □

Proposition 3.2. *For any $\delta = (\alpha, \beta, \gamma)$ and any $n \in \mathbb{N}$,*

$$\mathcal{M}^\beta(n) \leqslant \mathcal{N}^\delta([1,n]) + 1.$$

Proof. In the two-row graph $T^\delta([1,n])$ partition $R_1([1,n])$ into $R_2([1,n])$-similar equivalence classes $\{W_i\}$ so that two vertices $v_{x,1}$ and $v_{y,1}$ are in the same set W_i if they have the same neighbourhood in $R_2([1,n])$. By definition the number of such sets is $\mu(T^\delta([1,n]), R_1([1,n])) = \mathcal{N}^\delta([1,n])$. For $m < n$ partition $[1,m]$ into s sets $\{P_i\}$ such that $P_i = \{j : v_{j,1} \in W_i\}$. Then s is no more than the number of sets in $\{W_i\}$ by definition, but no less than $\mu(B^\beta([1,n]), [1,m]) - 1$, the number of equivalence classes that are $[m+1,n]$-similar (excluding, possibly, vertex m). This holds for all $m < n$, so

$$\begin{aligned}\mathcal{M}^\beta(n) - 1 &= \sup_{m<n} \mu(B^\beta([1,n]), [1,m]) - 1 \\ &\leqslant \mu(T^\delta([1,n]), R_1([1,n])) = \mathcal{N}^\delta([1,n]).\end{aligned}$$

□

3.2.2 Veins and Slices

We will start by considering only graph classes $\mathcal{G}^\delta$ for $\delta = (\alpha, \beta, \gamma)$ in which α is an infinite word from the alphabet $\{0,2\}$ and then in Subsection 3.2.4 we extend to the case where α is an infinite word from the alphabet $\{0,1,2,3\}$ by using the ideas of similarity, which apply equally to bipartite complements.

For $\mathcal{C}$ a proper hereditary subclass of $\mathcal{G}^\delta$, consider a specific embedding of a graph $G = (V,E) \in \mathcal{C}$ in $\mathcal{P}^\delta$, and recall that the induced subgraph of G on the vertices $V \cap C_{[j,j+k-1]}$

is denoted $G_{[j,j+k-1]}$. We are looking for a way to decompose the vertices of $G_{[j,j+k-1]}$ into a bounded number of subsets we term *veins* and *slices*, so that we can use this partition to create a clique-width expression for G using a bounded number of labels.

Let α be an infinite word over the alphabet $\{0,2\}$. A *vein* $\mathcal{V}$ of $G_{[j,j+k-1]}$ of length $t \leqslant k$ is a set of vertices $\{v_s, \dots, v_{s+t-1}\}$ in consecutive columns such that $v_y \in V \cap C_y$ for each $y \in \{s, \dots, s+t-1\}$ and for which $v_y \sim v_{y+1}$ for all $y \in \{s, \dots, s+t-2\}$.

We will call a vein of length $t = k$ a *full* vein and a vein of length $t < k$ a *part* vein. Note that as α comes from the alphabet $\{0,2\}$, for a vein $\{v_s, \dots, v_{s+t-1}\}$, v_{y+1} is no higher than v_y for each $y \in \{s, \dots, s+t-2\}$. A horizontal row of k vertices in $G_{[j,j+k-1]}$ is a full vein.

As G is Free($H^{\delta}_{j,1}(k,k)$) we know that no set of vertices of G induces $H^{\delta}_{j,1}(k,k)$. We consider this in terms of disjoint full veins of $G_{[j,j+k-1]}$. Note that k rows of vertices between column j and column $j+k-1$ are a set of k disjoint full veins and induce a graph isomorphic to $H^{\delta}_{j,1}(k,k)$. There are other sets of k disjoint full veins that form a graph isomorphic to $H^{\delta}_{j,1}(k,k)$, but some sets of k full veins do not. We wish to place a bound on the number of full veins and to do this our first task is to clarify when a set of k full veins has this property.

Let $\{v_j, \dots, v_{j+k-1}\}$ be a full vein such that each vertex v_x has coordinates (x, u_x) in $\mathcal{P}$, observing that $u_{x+1} \leqslant u_x$ for $x \in [j, j+k-2]$. We construct an *upper border* to be a set of vertical coordinates $\{w_j, \dots, w_{j+k-1}\}$ using the following procedure:

(1) Set $w_j = u_j$,

(2) Set $x = j+1$,

(3) if $\alpha_{x-1} = 2$ set $w_x = u_{x-1}$,

(4) if $\alpha_{x-1} = 0$ set $w_x = w_{x-1}$,

(5) set $x = x+1$,

(6) if $x = j+k$ terminate the procedure, otherwise return to step (3).

Given a full vein $\mathcal{V} = \{v_j, \dots, v_{j+k-1}\}$, define the *fat vein* $\mathcal{V}^f = \{v_{x,y} \in V(G_{[j,j+k-1]}) : x \in [j, j+k-1], y \in [u_x, w_x]\}$ (See examples shown in Figure 3.2).

Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be two full veins. Then we say they are *independent* if $\mathcal{V}^f_1 \cap \mathcal{V}^f_2 = \emptyset$ i.e. their corresponding fat veins are disjoint.

Proposition 3.3. $G_{[j,j+k-1]}$ *cannot contain more than* $(k-1)$ *independent full veins.*

Proof. We claim that k independent full veins $\{\mathcal{V}_1, \ldots, \mathcal{V}_k\}$ induce the forbidden graph $H^{\delta}_{j,1}(k,k)$.

Remembering $v_{x,y}$ is the vertex in the grid $\mathcal{P}$ in the x-th column and y-th row, let $w_{x,y}$ be the vertex in the y-th full vein $\mathcal{V}_y$ in column x. We claim the mapping $\phi(w_{x,y}) \to v_{x,y}$ is an isomorphism.

Consider vertices $w_{x,y} \in \mathcal{V}_y$ and $w_{s,t} \in \mathcal{V}_t$ for $t \geqslant y$. Then

(a) If $t = y$ (i.e the vertices are on the same vein) then both $w_{x,y} \sim w_{s,t}$ and $v_{x,y} \sim v_{s,t}$ if and only if $|x - s| = 1$ or $(x, s) \in \beta$,

(b) If $t > y$ and $x = s$ then both $w_{x,y} \sim w_{s,t}$ and $v_{x,y} \sim v_{s,t}$ if and only if $\gamma_x = 1$,

(c) If $t > y$ and $s = x + 1$ then both $w_{x,y} \not\sim w_{s,t}$ and $v_{x,y} \not\sim v_{s,t}$,

(d) If $t > y$ and $s = x - 1$ then both $w_{x,y} \sim w_{s,t}$ and $v_{x,y} \sim v_{s,t}$ if and only if $\alpha_s = 2$,

(e) If $t > y$ and $|s-x| > 1$ then both $w_{x,y} \sim w_{s,t}$ and $v_{x,y} \sim v_{s,t}$ if and only if $(x, s) \in \beta$.

Hence, $w_{x,y} \sim w_{s,t}$ if and only if $v_{x,y} \sim v_{s,t}$ and ϕ is an isomorphism from k independent full veins to $H^{\delta}_{j,1}(k,k)$. □

3.2.3 Vertex colouring

Our objective is to identify conditions on (recurrent) $\delta \in \Delta$ that make $\mathcal{G}^{\delta}$ a minimal class of unbounded clique-width. For such a δ it is sufficient to show that any graph G in a proper hereditary subclass $\mathcal{C}$ has bounded linear clique-width. In order to do this we will partition G into manageable sections (which we will call "panels"), the divisions between the panels chosen so that they can be built separately and then 'stuck' back together again, using a linear clique-width expression requiring only a bounded number of labels. In this section we describe a vertex colouring that will lead (in Section 3.2.5) to the construction of these panels.

As previously noted, for any subclass $\mathcal{C}$ there exist j and k such that $\mathcal{C} \subseteq \text{Free}(H^{\delta}_{j,1}(k,k))$. As δ is recurrent, if we let $\delta^* = \delta_{[j,j+k-1]}$ be the k-factor that defines the forbidden graph $H^{\delta}_{j,1}(k,k)$, we can find δ^* in δ infinitely often, and we will use these instances of δ^* to divide our embedded graph G into the required panels.

Firstly, we will construct a maximal set $\mathbb{B}$ of independent full veins for $G_{[j,j+k-1]}$, a section of G that by Proposition 3.3 cannot have more than $(k-1)$ independent full veins. For any two such veins, $\mathcal{V}_1$ and $\mathcal{V}_2$, we will say $\mathcal{V}_1$ is 'below' $\mathcal{V}_2$ where in each column of $G_{[j,j+k-1]}$ the vertex in $\mathcal{V}_1$ has a lower y coordinate than the vertex in $\mathcal{V}_2$ (remembering that the rows of the grid $\mathcal{P}$ are indexed from the bottom). We start with the lowest full vein and then keep adding the next lowest independent full vein until the process is exhausted.

Note that the next lowest independent full vein is unique because if we have two full veins $\mathcal{V}_1, \mathcal{V}_2$ with vertices $\{v_j, \ldots, v_{j+k-1}\}$ and $\{v'_j, \ldots, v'_{j+k-1}\}$ respectively then they can be combined to give $\{min(v_j, v'_j), \ldots, min(v_{j+k-1}, v'_{j+k-1})\}$ which is a full vein with a vertex in each column at least as low as the vertices of $\mathcal{V}_1$ and $\mathcal{V}_2$.

Let $\mathbb{B}$ contain $b < k$ independent full veins, numbered from the bottom as $\mathcal{V}_1, \cdots, \mathcal{V}_b$ such that any other full vein not in $\mathbb{B}$ must have a vertex in common with a fat vein $\mathcal{V}^f_y$ corresponding to one of the veins $\mathcal{V}_y$ of $\mathbb{B}$.

Let $u_{x,y}$ be the lowest vertical coordinate and $w_{x,y}$ the highest vertical coordinate of vertices in $\mathcal{V}^f_y \cap C_x$. We define $\mathcal{S}_0 = \{v_{x,y} \in V(G_{[j,j+k-1]}) : x \in [j, j+k-1], y < u_{x,1}\}$, $\mathcal{S}_b = \{v_{x,y} \in V(G_{[j,j+k-1]}) : x \in [j, j+k-1], y > w_{x,b}\}$, and for $y = 1, \ldots, b-1$ we define:

$$\mathcal{S}_i = \{v_{x,y} \in V(G_{[j,j+k-1]}) : x \in [j, j+k-1], w_{x,i} < y < u_{x,i+1}\}$$

This gives us $b+1$ *slices* $\{\mathcal{S}_0, \mathcal{S}_1, \cdots, \mathcal{S}_b\}$.

We partition the vertices in the fat veins and the slices into sets which have similar neighbourhoods, which will facilitate the division of G into panels. We colour the vertices of $G_{[j,j+k-1]}$ so that each slice has green/pink vertices to the left and red vertices to the right of the partition, and each fat vein has blue vertices (if any) to the left and yellow vertices to the right. Examples of vertex colourings are shown in Figure 3.2.

Colour the vertices of each slice $\mathcal{S}_i$ as follows:

- Colour any vertices in the left-hand column green. Now colour green any remaining vertices in the slice that are connected to one of the green left-hand column vertices by a part vein that does not have a vertex in common with any of the fat veins corresponding to the full veins in $\mathbb{B}$.
- Locate the column t of the right-most green vertex in the slice. If there are no green vertices set $t = s = j$. If $t > j$ then choose s in the range $j \leqslant s < t$ such that s is the highest column index for which $\alpha_s = 2$. If there are no columns before t for which $\alpha_s = 2$ then set $s = j$. Colour pink any vertices in the slice (not already coloured) in columns j to s which are below a vertex already coloured green.

- Colour any remaining vertices in the slice red.

Note that no vertex in the right-hand column can be green because if there was such a vertex then this would contradict the fact that there can be no full veins other than those which have a vertex in common with one of the fat veins corresponding to the full veins in $\mathbb{B}$. Furthermore, no vertex in the right hand column can be pink as this would contradict the fact that every pink vertex must lie below a green vertex in the same slice.

Colour the vertices of each fat vein $\mathcal{V}_i^f$ as follows:

- Let s be the column as defined above for the slice immediately above the fat vein. If $s = j$ colour the whole fat vein yellow. If $s > j$ colour vertices of the fat vein in columns j to s blue and the rest of the vertices in the fat vein yellow.

When we create a clique-width expression we will be particularly interested in the edges between the blue and green/pink vertices to the left and the red and yellow vertices to the right.

Proposition 3.4. *Let v be a red vertex in column x and slice $\mathcal{S}_i$.*

If u is a blue, green or pink vertex in column $x-1$ then

$$uv \in E(G) \textit{ if and only if } \alpha_{x-1} = 2 \textit{ and } u \in \mathcal{V}_{i+1}^f \cup \mathcal{S}_{i+1} \cup \cdots \cup \mathcal{V}_b^f \cup \mathcal{S}_b.$$

Similarly, if u is a blue, green or pink vertex in column $x+1$ then

$$uv \in E(G) \textit{ if and only if } \alpha_x = 2 \textit{ and } u \in \mathcal{S}_0 \cup \mathcal{V}_1^f \cup \mathcal{S}_1 \cup \cdots \cup \mathcal{V}_i^f \cup \mathcal{S}_i.$$

Proof. Note that as u and v are in consecutive columns we need only consider α-edges.

If u is green in column $x-1$ of $\mathcal{S}_i$ then red v in column x of $\mathcal{S}_i$ cannot be adjacent to u as this would place red v on a green part-vein which is a contradiction. Likewise, if u is green in column $x+1$ of $\mathcal{S}_i$ then red v in column x of $\mathcal{S}_i$ must be adjacent to u since if it was not adjacent to such a green vertex in the same slice then this implies the existence of a green vertex above the red vertex in the same column which contradicts the colouring rule to colour pink any vertex in columns j to s below a vertex coloured green.

The other adjacencies are straightforward. □

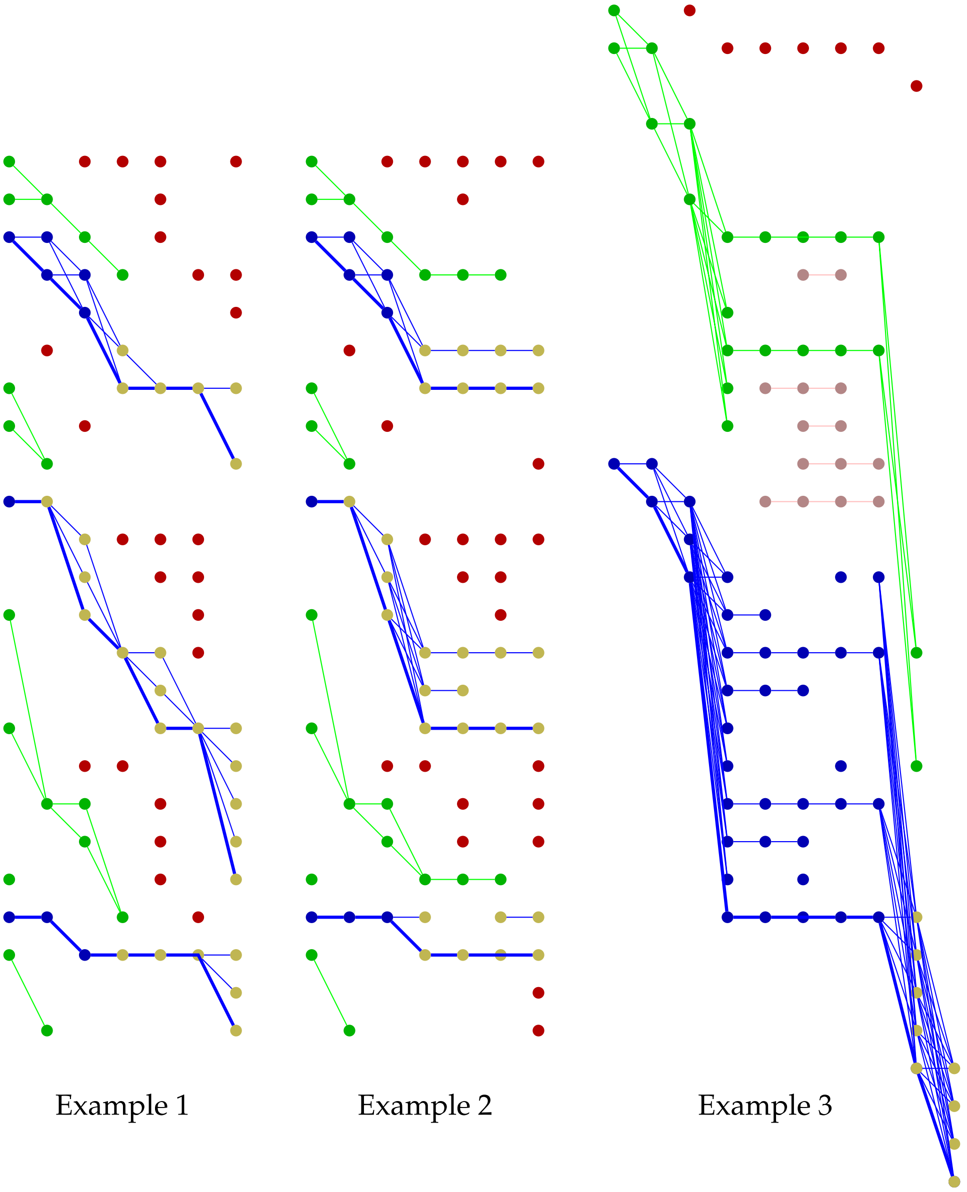


Figure 3.2: Examples of vein and slice colouring – a 222222, a 222000 and a 222000022 factor, with vertices coloured blue, green, pink, red and yellow as described. The only edges shown are the veins (bold blue), other edges in the fat veins (blue), part veins that start on the left column but do not reach the right column (green) and related pink rows.

Proposition 3.5. *Let v be a yellow vertex in column x and fat vein $\mathcal{V}_i^f$.*

If u is a blue, green or pink vertex in column $x-1$ then

$$uv \in E(G) \text{ if and only if } \alpha_{x-1} = 2 \text{ and } u \in \mathcal{V}_i^f \cup \mathcal{S}_i \cup \cdots \cup \mathcal{V}_b^f \cup \mathcal{S}_b.$$

Similarly, if u is a blue, green or pink vertex in column $x+1$ then

$$uv \in E(G) \text{ if and only if } \alpha_x = 2 \text{ and } u \in \mathcal{S}_0 \cup \mathcal{V}_1^f \cup \mathcal{S}_1 \cup \cdots \cup \mathcal{V}_{i-1}^f \cup \mathcal{S}_{i-1}.$$

Proof. Note that as u and v are in consecutive columns we need only consider α-edges.

If u is blue in column $x-1$ of $\mathcal{V}_i^f$ then yellow v in column x of $\mathcal{V}_i^f$ must be adjacent to u from the definition of a fat vein. Equally, from the colouring definition for a fat vein there cannot be a blue vertex in column $x+1$ of $\mathcal{V}_i^f$ if there is a yellow vertex in column x of $\mathcal{V}_i^f$.

The other adjacencies are straightforward. □

Having established these propositions, as the pink and green vertices in a particular slice and column have the same adjacencies to the red and yellow vertices, we now combine the green and pink sets and simply refer to them all as *green*.

3.2.4 Extending α to the 4-letter alphabet

Our analysis so far has been based on α being a word from the alphabet $\{0, 2\}$. We now use the following lemma to extend our colouring to the case where α is a word over the 4-letter alphabet $\{0, 1, 2, 3\}$.

Let α be an infinite word over the alphabet $\{0, 1, 2, 3\}$ and α^+ be the infinite word over the alphabet $\{0, 2\}$ such that for each $x \in \mathbb{N}$,

$$\alpha_x^+ = \begin{cases} 0 & \text{if } \alpha_x = 0 \text{ or } 1, \\ 2 & \text{if } \alpha_x = 2 \text{ or } 3. \end{cases}$$

Denoting $\delta = (\alpha, \beta, \gamma)$ and $\delta^+ = (\alpha^+, \beta, \gamma)$, let $G = (V, E)$ be a graph in the class $\mathcal{G}^\delta$ with a particular embedding in the vertex grid $V(\mathcal{P})$. We will refer to $G^+ = (V, E^+)$ as the graph with the same vertex set V as G from the class $\mathcal{G}^{\delta^+}$.

Lemma 3.6. *For any subset of vertices $U \subseteq V$, 2 vertices of U in the same column of $V(\mathcal{P})$ are $V \setminus U$-similar in G if and only if they are $V \setminus U$-similar in G^+.*

Proof. Let u_1 and u_2 be two vertices in U in the same column x and v be a vertex of $V \setminus U$ in column y. If $x = y$ then v is in the same column as u_1 and u_2 and is either adjacent to both or neither depending on whether there is a γ-clique on column x, which is the same in both G and G^+. If $|x - y| > 1$ then v is adjacent to both u_1 and u_2 if and only if there is a bond (x, y) in β, which is the same in both G and G^+.

If $y = x + 1$ then the adjacency of v to u_1 and u_2 is determined by α_x in G and α_x^+ in G^+. If $\alpha_x = \alpha_x^+$ (i.e. both 0 or both 2) then the adjacencies are the same in G and G^+. If $\alpha_x = 1$ and $\alpha_x^+ = 0$, then u_1 and u_2 are both adjacent to v in G if and only if they are both non-adjacent to v in G^+. If $\alpha_x = 3$ and $\alpha_x^+ = 2$, then u_1 and u_2 are both adjacent to v in G if and only if they are both non-adjacent to v in G^+.

Hence u_1 and u_2 have the same neighbourhood in $V \setminus U$ in G if and only if they have the same neighbourhood in $V \setminus U$ in G^+. □

Lemma 3.7. *For a graph* $G \in \mathcal{G}^\delta \cap \text{Free}(H^\delta_{j,1}(k,k))$ *and* G^+ *defined as above, let the vertices of* $G^+_{[j,j+k-1]}$ *be coloured as per Section 3.2.3. Then the same colouring applied to the vertices of* $G_{[j,j+k-1]}$ *has the property that a column of* $G_{[j,j+k-1]}$ *can be partitioned into at most* $k-1$ *disjoint blue sets and* k *disjoint green sets, so that any red or yellow vertex is either adjacent to all or none of a given green/blue vertex set.*

Proof. As α^+ is a word over the alphabet $\{0, 2\}$ the results of Sections 3.2.2 and 3.2.3 can be applied, in particular Propositions 3.3, 3.4 and 3.5. It follows that for $G^+_{[j,j+k-1]}$:

- there are no more than $(k-1)$ independent full veins, and consequently at most k slices,
- two blue vertices in the same fat vein and column have the same red/yellow neighbourhood, and
- two green vertices in the same slice and column have the same red/yellow neighbourhood.

Lemma 3.6, with U^b and U^g being the blue and green vertices respectively, and $U = U^b \cup U^g$, tells us that these statements also apply to $G_{[j,j+k-1]}$ and the result follows. □

3.2.5 Panel construction

We have, from the previous subsections, a decomposition of $G_{[j,j+k-1]}$ into veins and slices, and we want to use these to build up sections of G one at a time, where the interactions between each section are controlled; we will call these sections of G "panels".

To recap, $\delta^* = \delta_{[j,j+k-1]}$ is the k-factor that defines the forbidden graph $H^{\delta}_{j,1}(k,k)$ and we will use the repeated instances of δ^* to divide our embedded graph G into panels.

Define $t_0, t_1, \dots, t_z$ where t_0 is the index of the column before the first column of the embedding of G, t_z is the index of the last column of the embedding of G and t_i $(0 < i < z)$ represents the rightmost letter index of the i-th copy of δ^* in δ , such that $t_i > k+t_{i-1}$ to ensure the copies are disjoint. Hence, the i-th disjoint copy of δ^* in δ corresponds to columns $C_{[t_i-k+1,t_i]}$ of $\mathcal{P}^{\delta}$ and we denote the induced graph on these columns $G_i = G_{[t_i-k+1,t_i]}$ and denote G_i^+ as the corresponding graph in G^+.

Colour the vertices of G_i^+ blue, yellow, green or red as described in Section 3.2.3 and then apply the same colouring to the vertices of G_i. Call these G_i vertex sets U_i^b, U_i^y, U_i^g and U_i^r respectively. Denote U_1^w as the vertices in $G_{[t_0+1,t_1-k]}$, and for $1 < i < z$ denote U_i^w the set of vertices in $G_{[t_i+1,t_{i+1}-k]}$ and colour the vertices in each U_i^w white.

We now create a sequence of *panels*, the first panel is $P_1 = U_1^w \cup U_1^g \cup U_1^b$, and subsequent panels given by

$$P_i = U_{i-1}^y \cup U_{i-1}^r \cup U_i^w \cup U_i^g \cup U_i^b.$$

These panels create a disjoint partition of the vertices of our embedding of G. The following lemma will be used to put a bound on the number of labels required in a linear clique-width expression to create edges between panels. We denote $\mathbb{P}_i = \cup_{s=1}^{i} P_s$.

Lemma 3.8. *Let (α, β, γ) be a recurrent δ-triple where α is an infinite word over the alphabet $\{0,1,2,3\}$, γ is an infinite binary word and β is a bond set which has bounded $\mathcal{M}^{\beta}$, so that $\mathcal{M}^{\beta}(n) < M$ for all $n \in \mathbb{N}$.*

Then for any graph $G = (V,E) \in \mathcal{G}^{\delta} \cap \mathrm{Free}(H^{\delta}_{j,1}(k,k))$ for some $j,k \in \mathbb{N}$ with vertices V partitioned into panels $\{P_1, \dots, P_z\}$ and $1 \leqslant i \leqslant z$,

$$\mu(G, V \setminus \mathbb{P}_i) < M + 2k^2.$$

Proof. Considering the three sets of vertices $\mathbb{P}_i \setminus (U_i^b \cup U_i^g)$, U_i^b and U_i^g in graph G separately, we have:

(a) the number of distinct neighbourhoods of the vertex set $V \setminus \mathbb{P}_i$ in the vertex set $\mathbb{P}_i \setminus (U_i^b \cup U_i^g)$ is bounded by M.

(b) the number of distinct neighbourhoods of the vertex set $V \setminus \mathbb{P}_i$ in the vertex set U_i^b is bounded by $k(k-1)$, noticing that from Lemma 3.7 two blue vertices in the same fat vein and column have the same neighbourhood in $V \setminus \mathbb{P}_i$.

(c) the number of distinct neighbourhoods of the vertex set $V \setminus \mathbb{P}_i$ in the vertex set U_i^g is bounded by $k(k-1)$, noticing that from Lemma 3.7 two green vertices in the same slice and column have the same neighbourhood in $V \setminus \mathbb{P}_i$.

This covers all vertices of $\mathbb{P}_i$ so

$$\mu(G, V \setminus \mathbb{P}_i) \leqslant M + k(k-1) + k(k-1) < M + 2k^2.$$

□

3.2.6 When $\mathcal{G}^\delta$ is a minimal class of unbounded clique-width

At the beginning of Section 3.2 we stated that our strategy for proving that an arbitrary graph G in a proper hereditary subclass of $\mathcal{G}^\delta$ has bounded linear clique-width (and hence bounded clique-width) is to define an algorithm to create a linear clique-width expression that allows us to recycle labels so that we can put a bound on the total number of labels required, however many vertices there are in G.

We do this by constructing a linear clique-width expression for each panel P_i in G in a linear sequence, leaving the labels on each vertex of previously constructed panels $\mathbb{P}_{i-1}$ with an appropriate label to allow edges to be constructed between the current panel/vertex and previous panels. To be able to achieve this we require the following ingredients:

(a) δ to be recurrent so we can create the panels,

(b) a bound on the number of labels required to create each new panel,

(c) a process of relabelling so that we can leave appropriate labels on each vertex of the current panel to enable connecting to previous panels, before moving on to the next panel, and

(d) a bound on the number of labels required to create edges to previously constructed panels.

We have (a) by assumption and we will deal with (c) and (d) in the proof of Theorem 3.11. The next two lemmas show how we can restrict δ further, using a new concept of 'gap factors', to ensure (b) is achieved.

Lemma 3.9. *For any δ and graph $G \in \mathcal{G}^\delta$ and any $j_1, j_2 \in \mathbb{N}$ where $|j_2 - j_1| = \ell - 1$*

$$lcw(G_{[j_1, j_2]}) \leqslant 2\ell$$

Proof. We construct $G_{[j_1,j_2]}$ using a row-by-row linear method, starting in the bottom left. For each of the ℓ columns, we create 2 labels: one label $c_1, \dots, c_\ell$ for the vertex in the *current* row being constructed, and one label $e_1, \dots, e_\ell$ for the vertices in all *earlier* rows.

For the first row, we insert the (max) ℓ vertices using the labels $c_1, \dots, c_\ell$, and since every vertex has its own label we can insert all necessary edges. Now relabel $c_i \to e_i$ for each i.

Suppose that the first r rows have been constructed, in such a way that every existing vertex in column i has label e_i. We insert the (max) ℓ vertices in row $r+1$ using labels $c_1, \dots, c_\ell$. As before, every vertex in this row has its own label, so we can insert all edges between vertices within this row. Next, note that any vertex in this row has the same relationship with all vertices in rows $1, \dots, r$ of any column i. Since these vertices all have label e_i and the vertex in row $r+1$ has its own label, we can add edges as determined by α, β and γ as necessary. Finally, relabel $c_i \to e_i$ for each i, move to the next row and repeat until all rows have been constructed. □

We will call a factor of a δ-triple between, and including, some consecutive disjoint pair of occurrences of a k-factor $\delta^* = \delta_{[j,j+k-1]}$, a δ^**-gap factor*. An $\mathcal{N}^\delta$*-bounded recurrent* δ-triple is a recurrent triple where, for any factor δ^* and any δ^*-gap factor δ_Q, the value of $\mathcal{N}^\delta(Q)$ is bounded by a function of δ^* only (i.e. it is bounded irrespective of the δ^*-gap factor chosen). In particular, from Lemma 2.14, it follows that if δ is $\mathcal{N}^\delta$-bounded recurrent then there is a bound on the number of 2s and 3s in the α component of any δ^*-gap factor.

If δ is almost periodic, so that for any factor δ^* of δ every factor of δ of length at least $\mathcal{L}(\delta^*)$ contains δ^*, then each δ^*-gap factor δ_Q covers a maximum of $\mathcal{L}(\delta^*)+k$ columns. As a consequence of Lemma 3.9, $\mathcal{N}^\delta(Q)$ is bounded by $2(\mathcal{L}(\delta^*)+k)$ (i.e a function of δ^* only) irrespective of the δ^*-gap factor chosen. Hence, every almost periodic δ-triple is also $\mathcal{N}^\delta$-bounded recurrent.

In addition, we know there exist $\mathcal{N}^\delta$-bounded recurrent δ-triples which are not almost periodic – see Section 3.4.

Lemma 3.10. *Let δ be an $\mathcal{N}^\delta$-bounded recurrent triple with k-factor $\delta^* = \delta_{[j,j+k-1]}$. Then for any graph $G \in \mathcal{G}^\delta$, where $V[G] \subseteq C_Q$ where Q is an interval such that δ_Q is a factor of a δ^*-gap factor, there exists a bound on the linear clique-width of G that is a function of δ^* only.*

Proof. As δ is an $\mathcal{N}^\delta$-bounded recurrent triple there exists a bound $N(\delta^*)$ on $\mathcal{N}^\delta(Q)$, where Q is any interval such that δ_Q is a subset of a δ^*-gap factor. It follows from

Lemma 2.14 that there is a bound, say $J(\delta^*)$, on the number of 2s and 3s in the α factor of any δ^*-gap factor δ_Q.

We can use the row-by-row linear method from the proof of Lemma 2.16 to show that for any graph $G \in \mathcal{G}^\delta$, with $V[G] \subseteq C_Q$ we have $lcw(G) \leqslant 2J + N + 2$. □

We are now in a position to define a set of hereditary graph classes $\mathcal{G}^\delta$ that are minimal of unbounded clique-width. We will denote $\Delta_{min} \subseteq \Delta$ as the set of all δ-triples in Δ with the characteristics:

(a) δ is $\mathcal{N}^\delta$-bounded recurrent, and

(b) the bond set β has bounded $\mathcal{M}^\beta$.

Theorem 3.11. *If $\delta \in \Delta_{min}$ then $\mathcal{G}^\delta$ is a minimal hereditary class of both unbounded linear clique-width and unbounded clique-width.*

Proof. $\mathcal{G}^\delta$ has unbounded clique-width since $\delta \in \Delta$. We now show that if $\delta \in \Delta_{min}$ then every proper hereditary subclass $\mathcal{C} \subsetneq \mathcal{G}^\delta$ has bounded linear clique-width. From the introduction to this section we know that for such a subclass $\mathcal{C}$ there must exist some $H^\delta_{j,1}(k,k)$ for some j and $k \in \mathbb{N}$ such that $\mathcal{C} \subseteq Free(H^\delta_{j,1}(k,k))$.

Using the same column indices $\{t_i\}$ used for panel construction of a graph $G \in \mathcal{G}^\delta$ in Section 3.2.5, let the i-th δ^*-gap factor be denoted δ_{q_i} where $q_1 = [t_0 + 1, t_1]$ and $q_i = [t_{i-1} - k + 1, t_i]$ for $1 < i < z$. Note that for every i, $P_i \subseteq C_{q_i}$. From Lemma 3.10 we know there exist J and $N \in \mathbb{N}$, each a function of δ^* only, such that the number of labels required to construct each panel P_i by the row-by-row linear method for all $i \in \mathbb{N}$ is no more than $2J + N + 2$.

As the bond-set β has bounded $\mathcal{M}^\beta$, let $M \in \mathbb{N}$ be a constant such that $\mathcal{M}^\beta(n) < M$ for all $n \in \mathbb{N}$.

Although a single panel P_i can be constructed using at most $2J+N+2$ labels, we need to be able to recycle labels so that we can construct any number of panels with a bounded number of labels. We will show that any graph $G \in Free(H^\delta_{j,1}(k,k))$ can be constructed by a linear clique-width expression that only requires a number of labels determined by the constants M, N, J and k.

For our construction of panel P_i, we will use the following set of $4k^2 + MN + M + 2J + 2$ labels:

- 2 *current vertex labels*: a_1 and a_2;

- J *current row labels*: $\{c_y : y = 1, \dots, J\}$ for first J columns;
- J *previous row labels*: $\{p_y : y = 1, \dots, J\}$ for first J columns;
- MN *partition labels*: $\{s_{x,y} : x = 1, \dots, M, y = 1, \dots, N\}$, for vertices in $U^y_{i-1} \cup U^r_{i-1} \cup U^w_i$;
- k^2 *blue current panel labels*: $\{bc_{x,y} : x = 1, \dots, k, y = 1, \dots, k\}$, for vertices $\mathcal{V}^f_{i,x} \cap U^b_i \cap C_y$;
- k^2 *blue previous panel labels*: $\{bp_{x,y} : x = 1, \dots, k, y = 1, \dots, k\}$, for vertices $\mathcal{V}^f_{i-1,x} \cap U^b_{i-1} \cap C_y$;
- k^2 *green current panel labels*: $\{gc_{x,y} : x = 0, \dots, k-1, y = 1, \dots, k\}$, for vertices $\mathcal{S}_{i,x} \cap U^g_i \cap C_y$;
- k^2 *green previous panel labels*: $\{gp_{x,y} : x = 0, \dots, k-1, y = 1, \dots, k\}$, for vertices $\mathcal{S}_{i-1,x} \cap U^g_{i-1} \cap C_y$;
- M *bond labels*: $\{m_y : y = 1, \dots, M\}$, for vertices in previous panels for creating the β-bond edges between columns.

We carry out the following iterative process to construct each panel P_i in turn.

Assume $\mathbb{P}_{i-1} = \cup_{s=1}^{i-1} P_s$ has already been constructed such that labels m_y, $bp_{x,y}$ and $gp_{x,y}$ have been assigned to the $M + 2k^2$ $V \setminus \mathbb{P}_{i-1}$-similar sets as described in Lemma 3.8.

Using the same column indices $\{t_i\}$ used for panel construction (Section 3.2.5) we assign a default partition label $s_{x,y}$ to each column of $U^y_{i-1} \cup U^r_{i-1} \cup U^w_i$ as follows:

(a) Consider the bond-graph $B^\beta([1, t_z])$ (Section 3.2.1). We partition the interval $Q = [t_{i-1} - k + 1, t_i - k]$ into $[t_i - k + 1, t_z]$-similar sets of which there are at most M, and use label index x to identify values in Q in the same $[t_i - k + 1, t_z]$-similar set. Consequently, vertices in two columns of $U^y_{i-1} \cup U^r_{i-1} \cup U^w_i$ that have the same default label x value have the same neighbourhood in $G_{[t_i - k + 1, t_z]}$ and hence are in the same $V \setminus \mathbb{P}_i$-similar set.

(b) Consider the two-row graph $T^\delta(Q)$ (Section 2.3.1). We partition vertices in $R_1(Q)$ into $R_2(Q)$-similar sets of which there are at most N. We create a corresponding partition of the interval Q such that $v_{x,1}$ and $v_{y,1}$ are in the same equivalence class of $R_1(Q)$ if and only if x and y are in the same partition set of Q. We now use label index y to identify values in the same partition set. Consequently, vertices in two columns of $U^y_{i-1} \cup U^r_{i-1} \cup U^w_i$ that have the same default label y value have the same neighbourhood within G_Q.

We construct each panel P_i in the row-by-row linear method used for the graph with a finite number of 2s and 3s with bounded $\mathcal{N}^\delta$ constructed in Lemma 2.16. The current vertex always has a unique label. Thus, for each row, we use labels $c_1, \ldots, c_J$ for vertices in the first J columns and then alternate a_1 and a_2 for the current and previous vertices for the remainder of the row.

For each new vertex in the current row we add edges as follows:

(a) Insert required edges to the $\mathcal{M}^\beta + 2k^2$ $V \setminus \mathbb{P}_{i-1}$-similar sets – see Lemma 3.8. This is possible because vertices within each of these sets are either all adjacent to the current vertex or none of them are.

(b) Insert required edges to vertices in the same or lower rows in the current panel. This is possible as these vertices all have labels p_y, $s_{x,y}$, $bc_{x,y}$ or $gc_{x,y}$ and, from the construction, vertices with the same y value are either all adjacent to the current vertex or none of them are.

Following completion of edges to the current vertex, we relabel the previous vertex as follows:

- from c_y to p_y if it is in the first J columns,
- from a_2 (or a_1) to its default partition label $s_{x,y}$ if it is in $U^y_{i-1} \cup U^r_{i-1} \cup U^w_i$ but not in the first J columns.
- from a_2 (or a_1) to $bc_{x,y}$ if it is in $\mathcal{V}^f_{i,x} \cap U^b_i$, and
- from a_2 (or a_1) to $gc_{x,y}$ if it is in $\mathcal{S}_{i,x} \cap U^g_i$.

We now repeat for the next row of panel P_i.

Once panel P_i is complete, relabel as follows:

Relabel vertices in accordance with their $V \setminus \mathbb{P}_i$-similar set, of which there are at most M. Note from Proposition 3.1, that two vertices with the same label m_y from the previous $\mathbb{P}_{i-1}$ partition sets will still need the same label in $\mathbb{P}_i$. Two equivalence classes from the $\mathbb{P}_{i-1}$ partition may merge to form a new equivalence class in the $\mathbb{P}_i$ partition. Hence, it is possible to relabel with the same label the old equivalence classes that merge, and then use the spare m_y labels for any new equivalence classes that appear. We never need more than M such labels.

Also relabel all vertices with labels $bp_{x,y}$, $gp_{x,y}$, p_y and $s_{x,y}$ with the relevant bond label m_y of their $V \setminus \mathbb{P}_i$-similar set. This is possible for the vertices labelled $s_{x,y}$ as the index x signifies their $V \setminus \mathbb{P}_i$-similar set.

Now relabel $bc_{x,y} \to bp_{x,y}$ and $gc_{x,y} \to gp_{x,y}$ ready for the next panel. For the next panel we can reuse labels a_1, a_2, c_y, p_y, $s_{x,y}$, $bc_{x,y}$ and $gc_{x,y}$ as necessary.

This process repeated for all panels completes the construction of G.

The maximum number of labels required to construct any graph $G \in \text{Free}(H^{\delta}_{j,1}(k,k))$ is $4k^2 + MN + M + 2J + 2$ and hence $\mathcal{C}$ has bounded linear clique-width. □

The conditions for δ to be in Δ_{min} are sufficient for the class $\mathcal{G}^{\delta}$ to be minimal. It is fairly easy to see that it is necessary for δ to be bounded recurrent. However, there remains a question regarding the necessity of the bond set β to have bounded $\mathcal{M}^{\beta}$. We have been unable to identify any $\delta \notin \Delta_{min}$ such that $\mathcal{G}^{\delta}$ is a minimal class of unbounded clique-width, hence:

Conjecture 3.12. *The hereditary graph class $\mathcal{G}^{\delta}$ is minimal of unbounded clique-width if and only if $\delta \in \Delta_{min}$.*

3.3 Uncountably many minimal classes

We now proceed to show that there is an uncountably infinite number of minimal classes with unbounded clique-width. To do this we will use the class of almost periodic sequences known as *Sturmian*. One definition of a Sturmian sequence is a binary sequence that has *complexity* $p_{\alpha}(n) = n+1$, where the complexity function $p_{\alpha}(n)$ is the number of different factors of length n in α [31].

An alternative characterisation of Sturmian sequences is as rotation sequences defined by an irrational number (see [31]), and hence it follows that the number of such sequences is uncountably infinite. We say that two sequences are *locally isomorphic* if they have the same factors. If two Sturmian sequences are locally isomorphic this means they have the same $n+1$ factors of length n out of a possible 2^n such factors [62]. Hence the set of Sturmian sequences with a particular set of factors is countable in number and so it follows there is an uncountable number of such sets with different factors.

For finite word α^* we denote *rev(α^*)* as the letters of α^* in reverse order (mirror image).

Lemma 3.13. *For $i = 1,2$ let $\delta^i = (\alpha^i, \emptyset, 0^{\infty})$ where α^i is an infinite binary word and let $\alpha^* = \alpha^2_{[a,a+k-2]}$ be a finite factor of α^2 for some $a \in \mathbb{N}$ of length $k-1$ ($k \geqslant 2$) with at least one 1. Further, let H^* be the graph $H^{\delta^2}_{a,1}(k,3) \in \mathcal{G}^{\delta^2}$.*

Then H^* can be embedded in $\mathcal{P}^{\delta^1}$ if and only if α^* or rev(α^*) is a factor of α_1.

Furthermore, if $k \geqslant 3$*, such embedding is only possible in* 3 *rows and* k *consecutive columns* $s, \ldots, s+k-1$ *when* $\alpha^1_{[s,s+k-2]} = \alpha^*$ *or rev(*α^**).*

Proof. Clearly, by its definition, H^* can be embedded in $\mathcal{P}^{\delta^1}$ in the way described if α^* or rev(α^*) is a factor of α^1. [To avoid much repetition in what follows we will just refer to α^* to mean α^* or rev(α^*).] We prove that H^* can only be embedded in $\mathcal{P}^{\delta^1}$ in the way described, and only if α^* is a factor of α_1, by induction on k.

Firstly, if $k = 2$ then $\alpha^* = 1$ and $H^* = C_6$, the cycle on 6 vertices. It is trivial to see that this can be embedded in $\mathcal{P}^{\delta^1}$ only if there is at least one 1 in α_1. In fact, C_6 can be embedded in two ways. Firstly, (Method 1) in the way described in the Lemma, with 6 vertices from 3 rows and 2 consecutive columns, or secondly, (Method 2) over 4 consecutive columns with 1 vertex from the first and last column and 2 from each of the middle columns.

If $k = 3$ then α^* must be 10, 01 or 11. H^* is a graph on 9 vertices as shown in Figure 3.3.

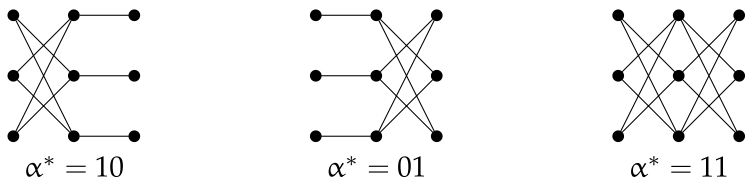


Figure 3.3: H^* when columns $k = 3$

Each version includes an induced subgraph C_6, with the addition of three more vertices. Method 2 no longer works as if the cycle C6 is spread over 4 columns then there are no possibilities for the remaining three vertices. Hence, H^* can only be embedded in $\mathcal{P}^{\delta^1}$ in the way described in the Lemma (Method 1) .

Next using the strong induction hypothesis, we assume that the Lemma is true for all words of length less than $k-1$. Thus if α^* contains a factor that is not a factor of α^1 then H^* cannot embed in $\mathcal{P}^{\delta^1}$.

If H^* does embed in $\mathcal{P}^{\delta^1}$ then, if α^{*-} is the word α^* without its last letter and H^{*-} the induced subgraph of H^* without its last column of vertices, we must have α^{*-} a factor of α^1 where H^{*-} can only embed in $\mathcal{P}^{\delta^1}$ by Method 1. Now it is straightforward to see that this cannot be extended to H^* if the next letter is not the same as the last letter of α^*, and that if it is the same, it can only be done by Method 1. □

Theorem 3.14. *There exists an uncountably infinite number of minimal hereditary classes of graphs of unbounded clique-width.*

Proof. There exists an uncountably infinite number of Sturmian binary sequences that are not locally isomorphic. Suppose we have Sturmian words α^1 and α^2 that have unique factors α^{*1} and α^{*2} with corresponding graphs H^{*1} and H^{*2} respectively. Then using Lemma 3.13, the class $\mathcal{G}^{\delta_1}$ does not contain the graph H^{*2} and the class $\mathcal{G}^{\delta_2}$ does not contain the graph H^{*1}. So $\mathcal{G}^{\delta_1}$ and $\mathcal{G}^{\delta_2}$ are different graph classes. It follows from Theorem 3.11 each one defines a different minimal hereditary class of graphs of unbounded clique-width. □

3.4 Classes defined by recurrent but not almost periodic words

In this section we identify examples of $\mathcal{N}^\delta$-bounded recurrent δ-triples that are not almost periodic. To do this we narrow our focus to consider only triples of the form $\delta = (\alpha, \emptyset, 0^\infty)$. Whilst we have identified an uncountable number of minimal classes based on almost periodic (Sturmian) words, it is more difficult to identify words α that are recurrent but not almost periodic, i.e. words in which each factor occurs infinitely many times, but where the gap between consecutive occurrences of a factor may be arbitrarily large.

One simple way to generate sequences is by substitution, and we use [31] as our reference work for this. For example, consider the infinite binary word ψ generated by an iterative substitution σ, beginning with 1 such that $\sigma(1) = 1010$ and $\sigma(0) = 0$.

If we denote $\sigma^n(1)$ as the n-th iteration beginning with $\sigma^0(1) = 1$ then

$$\sigma^n(1) = \sigma^{n-1}(\sigma(1)) = \sigma^{n-1}(1)\ 0\ \sigma^{n-1}(1)\ 0.$$

The first four iterates, and the start of ψ, are as follows.

$$\begin{aligned}
\sigma^1(1) &= 1010\\
\sigma^2(1) &= 1010\,0\,1010\,0\\
\sigma^3(1) &= 1010010100\,0\,1010010100\,0\\
\sigma^4(1) &= 1010010100010100101000\,0\,1010010100010100101000\,0\\
\psi &= 10100101000101001010000101001010001010010100000101001010 00\ldots
\end{aligned}$$

The word ψ has the following characteristics.

(i) The number of ones doubles with each iteration and therefore ψ contains an infinite number of ones.

(ii) ψ is a fixed point of σ (i.e. $\sigma(\psi) = \psi$).

(iii) By construction ψ is recurrent but is not almost periodic. The n^{th} iteration $\sigma^n(1)$ ends with a string of n zeros and hence ψ contains arbitrarily long strings of zeros. The gaps between non-zero factors must therefore widen as $n \to \infty$.

If $\delta = (\psi, \emptyset, 0^\infty)$ it is clear that $\mathcal{N}^\delta$ is unbounded since ψ has an infinite number of 1s so by Theorem 2.17 we have $\delta \in \Delta$. Furthermore, the bond set $\beta = \emptyset$ has bounded $\mathcal{M}^\beta$, so there is only one remaining condition to satisfy to show $\delta \in \Delta_{min}$.

We define the *weight* of a word α as the number of non-zero letters it has, which we will denote $|\alpha|_1$.

Lemma 3.15. $\delta = (\psi, \emptyset, 0^\infty)$ *is a* $\mathcal{N}^\delta$*-bounded recurrent* δ*-triple.*

Proof. We must show that for any k-factor $\delta^* = \delta_{[j,j+k-1]}$ and any δ^*-gap factor δ_Q, the value of $\mathcal{N}^\delta(Q)$ is bounded by a function of δ^* (i.e the quantity k) only. Since $\beta = \emptyset$ and $\gamma = 0^\infty$ then the only graph edges in $\mathcal{P}^\delta$ are defined by ψ, so we will just refer to ψ rather than δ in the following proof.

Suppose the longest subfactor of contiguous zeros in ψ^* is 0^k. It can be observed that $\sigma^n(1)$ ends with the factor 0^n. Hence ψ^* must have appeared by the $(k+1)$-th iteration, $\sigma^{k+1}(1)$ or it is not a factor of ψ. Since $|\sigma^{k+1}(1)|_1 = 2^{k+1}$, we have this as a bound on the weight between any consecutive occurrences. Thus, the value of $\mathcal{N}^\delta(Q)$ is bounded by 2^{k+1} for any δ^*-gap factor δ_Q, which is a function only of ψ^* (i.e. k the length of ψ, irrespective of the length of the interval Q). □

Theorem 3.16. $\delta = (\psi, \emptyset, 0^\infty) \in \Delta_{min}$.

Proof. We have shown that δ satisfies the conditions of Theorem 3.11. □

The paper by Collins et al [17] examined similar classes of graph restricted to $\delta = (\alpha, \emptyset, 0^\infty)$ with α over the alphabet $\{0, 1, 2\}$. They conjectured that such a class $\mathcal{G}^\delta$ is a minimal hereditary class of unbounded clique-width if and only if α is almost periodic and contains at least one 1 or 2. Thus, we have shown the 'if' statement to be true and provide a counterexample to contradict the only if statement.

We can extend this idea to construct other recurrent but not almost periodic triples $\delta^\rho = (\alpha^\rho, \emptyset, 0^\infty) \in \Delta_{min}$. Indeed, any iterative substitution σ_ρ where $\sigma_\rho(1) = \rho$ and $\sigma_\rho(0) = 0$ such that ρ is a finite binary word whose first letter is 1, last letter is 0, and with $|\rho|_1 \geqslant 2$ will define an infinite binary word α^ρ. For example if $\rho = 1100$ then the first four iterates are as follows:

$$\sigma^1_\rho(1) = 1100$$
$$\sigma^2_\rho(1) = 1100\,1100\,0\,0$$
$$\sigma^3_\rho(1) = 1100110000\,1100110000\,0\,0$$
$$\sigma^4_\rho 1) = 1100110000110011000000\,1100110000110011000000\,0\,0$$

Now, $|\sigma^n_\rho(1)|_1 = |\rho|^n_1$ and it follows using Lemma 3.15 that δ^ρ is a $\mathcal{N}^\delta$-bounded recurrent δ-triple, and hence $\delta^\rho \in \Delta_{min}$.

Finally, notice that our set of triples defining minimal hereditary graph classes of unbounded clique-width Δ_{min} does not include all recurrent δ triples. Indeed, for any infinite recurrent but not almost periodic binary word α, such that the triple $\delta = (\alpha, \emptyset, 0^\infty) \in \Delta_{min}$, then the triple $\overline{\delta} = (\overline{\alpha}, \emptyset, 0^\infty)$, where $\overline{\alpha}$ is the complement of α (i.e. inverting the 1s and 0s), is a recurrent δ-triple that is not $\mathcal{N}^\delta$-bounded recurrent and therefore does not lie in Δ_{min}.

3.5 Further examples of minimal classes from the framework

In Sections 3.3 and 3.4 we only used examples from the framework of triples $\delta = (\alpha, \emptyset, 0^\infty)$. However, the framework provides many other types of minimal classes when $\beta \neq \emptyset$ and $\gamma \neq 0^\infty$. Some examples of such $\delta = (\alpha, \beta, \gamma)$ values that yield a minimal class are shown in Table 3.1.

Example	α	β **($x, y \in \mathbb{N}$)**	γ	$\mathcal{M}^\beta$ bound
1.	0^∞	$\emptyset$	1^∞	1
2.	1^∞	$(1, x+2)$	0^∞	2
3.	$(23)^\infty$	$(x, x+2)$	0^∞	3
4.	0^∞	$(x, y) : \lvert x-y\rvert \neq 1, x-y \equiv 1 \pmod 2$	0^∞	3
5.	1^∞	$(x, y) : x \neq y, x-y \equiv 0 \pmod 2$	1^∞	2
6.	2^∞	$(x, y) : 1 < \lvert x-y\rvert \leqslant n$ (fixed n)	0^∞	n

Table 3.1: New minimal hereditary graph classes of unbounded clique-width

3.5.1 Final thoughts

The ideas of periodicity and recurrence are well established concepts when applied to symbolic sequences (i.e. words). Application to δ-triples and in particular β-bonds is rather different and needs further investigation.

The β-bonds have been defined as generally as possible, allowing a bond between any two non-consecutive columns. The purpose of this was to capture as many minimal classes in the framework as possible. However, it may be observed that the definition is so general that for any finite graph G it is possible to define β so that G is isomorphic to an induced subgraph of $B^{\beta}(Q)$ and hence $\mathcal{G}^{\delta}$.

In these $\mathcal{G}^{\delta}$ graph classes we have seen that unboundedness of clique-width is determined by the unboundedness of a parameter measuring the number of distinct neighbourhoods between two-rows. The minimal classes are those which satisfy defined recurrence characteristics and for which there is a bound on a parameter measuring the number of distinct neighbourhoods between vertices in one row.

Hence, whilst we have created a framework for many types of minimal classes, there may be further classes 'hidden' in the β-bonds.

Chapter 4

Sparse Classes and t-sails

4.1 Introduction

It has long been known that the following basic objects are obstructions to bounded tree-width: for arbitrarily large t, (1) the complete graph K_t, (2) the complete bipartite graph $K_{t,t}$, (3) a subdivision of the $(t \times t)$-wall and (4) the line graph of a subdivision of the $(t \times t)$-wall. We now add a further *boundary object* to this list, a t-*sail* (see Figure 4.1).

These objects have been discovered by studying sparse hereditary *path-star* classes, each of which consists of the finite induced subgraphs of a single infinite graph whose edges can be partitioned into a path (or forest of paths) with a forest of stars, characterised by an infinite word over a possibly infinite alphabet. We show that a path-star class whose infinite graph has an unbounded number of stars, each of which connects an unbounded number of times to the path, has unbounded tree-width. In addition, we show that such a class is not a subclass of the hereditary class of circle graphs.

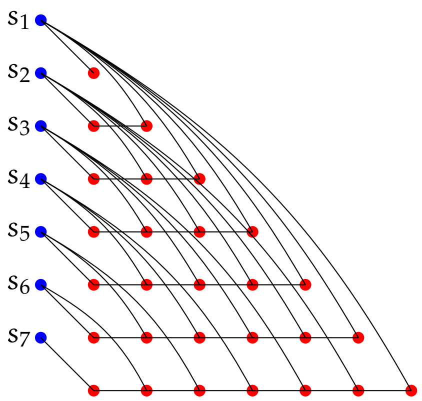


Figure 4.1: A 7-sail

We identify a collection of *nested* words with a recursive structure that exhibit interesting characteristics when used to define a path-star graph class. These graph classes do not contain any of the four basic obstructions but instead contain graphs that have large tree-width if and only if they contain arbitrarily large t-sails. We show that these classes are infinitely defined and, like classes of bounded degree or classes excluding a fixed minor, do not contain a minimal class of unbounded tree-width.

4.2 Nested Words

In this Section we introduce *nested words*, describe their characteristics and provide a range of examples and methods of generation to demonstrate that not only do they exist but are, in fact, quite common. These words are central to the analysis in Sections 4.4 and 4.5 where we show that path-star classes defined by such words have a number of interesting features: they are KKW-free (see Definition 1.28), t-sails are the basic objects obstructing bounded tree-width, they are infinitely defined and do not contain a minimal class of unbounded tree-width.

4.2.1 Definitions

Given a word α over an alphabet $\mathcal{A}$, and a sub-alphabet $\mathcal{S} \subset \mathcal{A}$, the *subword* of α restricted to $\mathcal{S}$ is the word derived from α by deleting all letters not in $\mathcal{S}$ and concatenating the remaining factors in the same order as they appear in α. We denote this subword as $\alpha^{\mathcal{S}}$.

We define *branched* and *nested* words as follows:

Definition 4.1. A word α over alphabet $\mathcal{A} = \mathbb{N}$ is *branched* if $\mathcal{A}$ can be partitioned into a finite *base set* $\mathcal{B}$ and ordered (by the order inherited from $\mathbb{N}$) *branch sets* $\{\mathcal{H}_1, \mathcal{H}_2, \dots\}$ which may be finite or infinite, such that:

- A base letter can appear after any other letter.
- The first letter in a branch $\mathcal{H}_i$ can only appear after a base letter.
- Any other letter in a branch $\mathcal{H}_i$ can only appear after the letter preceding it in the $\mathcal{H}_i$ order.

A maximal factor of branched word α containing no base letters is called a *branch* of α. Thus, a branch in α is preceded by a base letter unless it begins on the first letter of α

and is succeeded by a base letter unless it ends on the last letter of α. The letters in the branch come from one branch set, say $\mathcal{H}_i$, and appear, starting with the first letter in $\mathcal{H}_i$, in the defined order.

Definition 4.2. Let α be an infinite branched word over an infinite alphabet $\mathcal{A}$ with respect to base set $\mathcal{B}_\alpha$ and branch sets $\{\mathcal{H}_1^\alpha, \mathcal{H}_2^\alpha, \dots\}$ with the property that each letter in $\mathcal{A}$ appears an infinite number of times in α. Then α is $\mathfrak{b}$-*nested* if there exists a fixed positive integer $\mathfrak{b}$ such that any subset $\mathcal{S} \subseteq \mathcal{A}$ can be partitioned into a base set $\mathcal{B}_\mathcal{S}$ and ordered branch sets $\{\mathcal{H}_1^\mathcal{S}, \mathcal{H}_2^\mathcal{S}, \dots\}$, such that

- $|\mathcal{B}_\mathcal{S}| \leqslant \mathfrak{b}$,
- $\alpha^\mathcal{S}$ is a branched word with base set $\mathcal{B}_\mathcal{S}$, and
- the branch sets of $\alpha^\mathcal{S}$ are (possibly empty) subsets of the branch sets of α (i.e $\mathcal{H}_i^\mathcal{S} \subseteq \mathcal{H}_i^\alpha$ with the same ordering, for each i).

Example 4.3. *Consider the words σ^1, σ^2 and α where σ^1 is all* 1*s, σ^2 is the word*

$$232343234543234565432345676543234567 8\dots$$

and α is the word whose odd letters are σ^1 and even letters are σ^2, so that:

$$\alpha = 1\,2\,1\,3\,1\,2\,1\,3\,1\,4\,1\,3\,1\,2\,1\,3\,1\,4\,1\,5\,1\,4\,1\,3\,1\,2\,1\,3\,1\,4\,1\,5\,1\,6\dots$$

α is branched with $\mathcal{B} = \{1\}$ (in blue) and branch sets $\mathcal{H}_1 = \{2\}$, $\mathcal{H}_2 = \{3\}$, $\mathcal{H}_3 = \{4\}$ etc. It is an infinite word with an infinite alphabet $\mathcal{A} = \mathbb{N}$ where each letter appears an infinite number of times. However, it is not nested, since letting $\mathcal{S} = \mathcal{A} \setminus 1$ then $\alpha^\mathcal{S} = \sigma^2$ is not a branched word since there is no finite base set $\mathcal{B}_\mathcal{S}$ that can partition $\alpha^\mathcal{S}$ into one-letter branches that are subsets of the branch sets of α, which contradicts the third bullet-point in the definition of $\mathfrak{b}$-nested.

We refer generally to *nested words* when referring to a collection of $\mathfrak{b}$-nested words for some unspecified $\mathfrak{b}$. Some examples of nested words are given in Sections 4.2.2 and 4.2.3.

4.2.2 Examples : Arithmetic nested words

The following nested words have infinite branch sets.

- $\alpha(1)$: One branch set with increasing branch sizes (base letters shown in blue, gaps in word used to ease parsing):

$$\alpha(1) = 12\ 123\ 1234\ 12345\ 123456\ 1234567\ 12345678\ 123456789\ 1\ldots$$

- $\alpha(2)$: Two branch sets (residue classes mod 2) with increasing branch sizes:

$$\alpha(2) = 1\ 2\ 13\ 24\ 135\ 246\ 1357\ 2468\ 13579\ 2468\ 10\ldots$$

- $\alpha(k)$: k branch sets (residue classes mod k) with increasing branch sizes:

$$\alpha(k) = 1\ 2\ldots\ k\ \ 1\,(k+1)\ 2\,(k+2)\ldots\ k\,(2k)\ \ldots$$
$$1\,(k+1)\ (2k+1)\ 2\ (k+2)\ (2k+2)\ldots\ldots\ k\ (2k)\ (3k)\ \ldots$$

Lemma 4.4. *$\alpha(k)$ is a nested word for all $k \in \mathbb{N}$.*

Proof. Observe $\alpha(k)$ is branched with $\mathcal{A} = \mathbb{N}, \mathcal{B} = \{1, 2, \ldots, k\}, \mathcal{H}_1 = \{k+1, 2k+1, \ldots, nk+1, \ldots\}, \mathcal{H}_2 = \{k+2, 2k+2, \ldots, nk+2, \ldots\}, \ldots, \mathcal{H}_k = \{2k, 3k, \ldots, nk, \ldots\}$.

Let $\mathcal{S}$ be any subset of $\mathbb{N}$. Define $\mathcal{S}_i = \{n \in \mathcal{S} : n \equiv i \mod k\}$ for $1 \leqslant i \leqslant k$, $m_i = \min(\mathcal{S}_i)$ for $1 \leqslant i \leqslant k$, $\mathcal{B}_\mathcal{S} = \cup_{i=1}^{k} m_i$ and $\mathcal{H}_{\mathcal{S}i} = \mathcal{S}_i \setminus m_i$. Then $\alpha(k)^\mathcal{S}$ is branched with base $\mathcal{B}_\mathcal{S}$, $|\mathcal{B}_\mathcal{S}| \leqslant k$ and branch sets $\mathcal{H}_{\mathcal{S}i}$. Hence, $\alpha(k)$ satisfies the conditions of Definition 4.2. □

4.2.3 Examples: Power nested words

The following nested words have an infinite number of single letter branch sets.

4.2.3.1 q-ary representation

For natural numbers q and n, let n_q be the representation of n in q-ary. Let k be the number of trailing zeros of n_q and let j be the first non-zero digit from the right. Alternatively, there exist unique $j, k \in \mathbb{N}$ and $m \in \mathbb{N}_0$ such that $n = jq^k + mq^{k+1}$ $(1 \leqslant j \leqslant q-1)$.

We define infinite *power* words $\kappa(q)$ such that the n-th letter $\kappa(q)_n = i$ where $i = k(q-1) + j$ (e.g. for $q = 3$ we have $\kappa(3)_{45} = 6$ because 45 base 3 is 1200, so $k = 2, j = 2$ and $i = 2(3-1) + 2 = 6$).

This construction gives us:

$$\kappa(2) = 121\,3\,121\,4\,1213121\,5\,121312141213121\,6\ldots$$
$$\kappa(3) = 12\,3\,12\,4\,12\,5\,12312412\,6\,12312412\,7\,123124125\ldots$$
$$\kappa(4) = 123\,4\,123\,5\,123\,6\,123\,7\,123412351236123\,8\,1234\ldots$$
$$\kappa(5) = 1234\,5\,1234\,6\,1234\,7\,1234\,8\,1234\,9\,12345123461234 7\ldots$$

(Note that $\kappa(2)$ has previously appeared in the literature in the context of dense graphs [49] and in the context of sparse graphs in [11].)

These words can also be generated by a recurrence relation. For example, $\kappa(2)$ can be generated by the recurrence relation $\kappa(2) = \lim_{n\to\infty} \kappa(2)^n$ where $\kappa(2)^1 = 1$ and for $n > 1$, $\kappa(2)^n = \kappa(2)^{n-1}(n)\kappa(2)^{n-1}$. The first four iterates are as follows:

$$\kappa(2)^1 = 1, \kappa(2)^2 = 1\,2\,1, \kappa(2)^3 = 121\,3\,121, \kappa(2)^4 = 1213121\,4\,1213121$$

In general:

Proposition 4.5. *$\kappa(q)$ can be generated by the recurrence relation $\kappa(q) = \lim_{n\to\infty} \kappa(q)^n$ where $\kappa(q)^1 = 123\ldots\, q-1$ and for $n > 1$,*

$$\begin{aligned}\kappa(q)^n =&\kappa(q)^{n-1}((n-1)(q-1)+1)\kappa(q)^{n-1}((n-1)(q-1)+2)\ldots\\ &\kappa(q)^{n-1}((n-1)(q-1)+(q-2))\kappa(q)^{n-1}(n(q-1))\kappa(q)^{n-1}.\end{aligned}$$

Proof. Notice that by induction, $|\kappa(q)^1| = q-1$, $|\kappa(q)^2| = q^2-1$ and $|\kappa(q)^n| = q^n - 1$ and recall that for the x-th letter in $\kappa(q)$, where $x = jq^k + mq^{k+1}$, $\kappa(q)_x = k(q-1)+j$.

Therefore, in $\kappa(q)^n$ the letters in the interval $[jq^{n-1}+1, (j+1)q^{n-1}-1]$ for $0 \leqslant j \leqslant q-1$ are identical to $\kappa(q)^{n-1}$ and the letter in location jq^{n-1} for $0 \leqslant j \leqslant q-1$ is $(n-1)(q-1)+j$. The recurrence relation follows. □

Lemma 4.6. *$\kappa(q)$ is a nested word for all $q \in \mathbb{N}$.*

Proof. Observe $\kappa(q)$ is branched with $\mathcal{A} = \mathbb{N}, \mathcal{B} = \{1, 2, \ldots, q-1\}, \mathcal{H}_1 = \{q\}, \mathcal{H}_2 = \{q+1\}, \ldots, \mathcal{H}_k = \{q+k-1\}, \ldots$.

Let $\mathcal{S} = \{x_1, x_2, \ldots\}$ be a subset of $\mathbb{N}$, where $x_1 < x_2 \ldots$. Let n be a position in $\kappa(q)$ with the letter x_1 and let j, k, m be the unique integers such that $n = jq^k + mq^{k+1}$ and $x_1 = k(q-1)+j$ as described above. Further, let $x_i \geqslant x_1$ be the highest element of $\mathcal{S}$ such that $x_i < kq$. We claim that $\kappa(q)$ is nested over $\mathcal{S}$ with base $\mathcal{B}_{\mathcal{S}} = \{x_1, x_2, \ldots x_i\}$ of maximum size $q-1$ with single letter branch sets $\mathcal{H}^{\mathcal{S}}_1 = \{x_{i+1}\}, \ldots, \mathcal{H}^{\mathcal{S}}_t = \{x_{i+t}\}, \ldots$.

For any positive integer t the letter immediately preceding or succeeding an appearance of x_{i+t} in $\kappa(q)^{\mathcal{S}}$ is either x_1 or x_i. Hence, $\kappa(q)^{\mathcal{S}}$ is branched with base $\mathcal{B}_{\mathcal{S}}$, $|\mathcal{B}_{\mathcal{S}}| \leqslant q-1$ and branch sets $\mathcal{H}_t^{\mathcal{S}}$, and $\kappa(q)$ satisfies the conditions of Definition 4.2. □

4.2.3.2 Fibonacci representation

The well-known Fibonacci sequence of numbers is defined recursively as $F_0 = 0, F_1 = 1, F_n = F_{n-1} + F_{n-2}$ for $n \geqslant 2$ so that

$$F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, F_6 = 8, F_7 = 13, F_8 = 21, \ldots$$

The Fibonacci representation of a number is a sequence of 0s and 1s, rather like binary, except that a 1 in position k counting from the right represents F_{k+1} instead of 2^{k-1}. Note that to avoid having a repeated 1 in the sequence we start with F_2. So, for example, 101101 represents $2^5+2^3+2^2+2^0 = 32+8+4+1 = 45$ in binary but $F_7+F_5+F_4+F_2 = 13+5+3+1 = 22$ as a Fibonacci representation. In general, a Fibonacci representation is not unique. However, Zeckendorf [63] showed that every positive integer can be represented uniquely as the sum of non-consecutive Fibonacci numbers, so we will only use the Zeckendorf representation here.

We define the infinite word η such that $\eta_n = i$ if the first 1 in the Fibonacci representation of n (from the right) appears in position i (e.g. $n = 45$ which has Fibonacci representation 10010100 gives $\eta_{45} = 3$). Thus:

$$\eta = 12\,3\,1\,4\,12\,5\,1231\,6\,1231412\,7\,123141251231\,8\ldots$$

Also observe:

Proposition 4.7. *η can be generated by the recurrence relation $\eta = \lim_{n\to\infty} \eta^n$ where $\eta^1 = 1$, $\eta^2 = 12$ and for $n > 2$, $\eta^n = \eta^{n-1}(n)\eta^{n-2}$.*

Proof. Notice that by induction, $|\eta^1| = 1$, $|\eta^2| = 2$ and $|\eta^n| = F_{n+2} - 1$.

By definition, the first $F_{n+1} - 1$ letters of η^n must be η^{n-1}. Equally, the first 1 in the representation of F_{n+1} is in position n so the F_{n+1}-th letter of η^n is n. Also, the letters in the interval $[F_{n+1} + 1, F_{n+2} - 1]$ of length $F_n - 1$ are identical to the letters in the interval $[1, F_n - 1]$, and the recurrence relation follows. □

Lemma 4.8. *η is a nested word.*

Proof. Observe η is branched with $\mathcal{A} = \mathbb{N}, \mathcal{B} = \{1, 2\}, \mathcal{H}_1 = \{3\}, \mathcal{H}_2 = \{4\}, \ldots, \mathcal{H}_k = \{k+2\}, \ldots$.

Let $\mathcal{S} = \{x_1, x_2, \ldots\}$ be a subset of $\mathbb{N}$, where $x_1 < x_2 \ldots$, with $\mathcal{B}_{\mathcal{S}} = \{x_1, x_2\}$, $\mathcal{H}_1^{\mathcal{S}} = \{x_3\}, \ldots, \mathcal{H}_t^{\mathcal{S}} = \{x_{t+2}\}, \ldots$. Then, from the Zeckendorf representation, for any $t \geqslant 3$ the letter immediately preceding or succeeding an appearance of x_t in $\eta^{\mathcal{S}}$ is either x_1 or x_2. Hence, $\eta^{\mathcal{S}}$ is branched with base $\mathcal{B}_{\mathcal{S}}$, $|\mathcal{B}_{\mathcal{S}}| \leqslant 2$ and branch sets $\mathcal{H}_t^{\mathcal{S}}$, and η satisfies the conditions of Definition 4.2. □

4.2.4 Generating new nested words

Other nested words can be generated by similar recurrence relations to those given in Propositions 4.5 and 4.7, although not all such relations give nested words (for instance, if the recurrence relation does not result in a word over an infinite alphabet where each letter repeats an infinite number of times).

If α is a nested word and $\mathcal{L}$ a finite collection of letters (that may or may not be letters in $\mathcal{A}$, the alphabet of α) then inserting an arbitrary number of letters from $\mathcal{L}$ into arbitrary positions in α creates a new word that is also nested, since we can add the finite number of letters in $\mathcal{L}$ to the base $\mathcal{B}$. Hence, nested words are not rare.

Proposition 4.9. *There are uncountably many distinct nested words.*

Proof. Let α be a nested word and β an infinite (non-nested) binary word. Interlace the letters of α and β to create a new word γ, so that the even letters of γ are α and the odd letters β. γ is nested since we can just add the two letters of β to the base of α to create a finite base for γ. There are uncountably many distinct binary words which gives the result. □

However, the existence of a nested subword β in α is not sufficient to make α a nested word. The subword $\alpha \setminus \beta$ may have an unbounded base and contain elements that contradict our desired characteristics - see Section 4.4.1.

4.3 Path-star hereditary graph classes and t-sails

A convenient way to define a family of hereditary path-star classes as described in the introduction (see Definition 1.29) is to use an infinite word so that the i-th letter in the word indicates the star that connects to the i-th vertex in the path. We assume that all leaves of the stars embed in the path since non-embedding leaves have no effect on

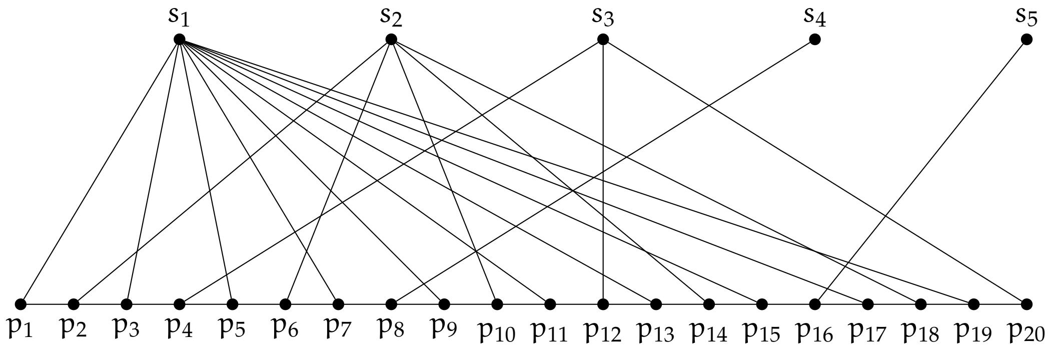


Figure 4.2: The first section of path-star graph $R^{\kappa(2)}$

tree-width, i.e., if G is a finite path-star graph and H is isomorphic to G except for the removal of all non-embedding leaves then $tw(H) = tw(G)$.

Let α be an infinite word over the alphabet $\mathcal{A} = \mathbb{N}$. We denote the path $P = (V_P, E_P)$ with vertices $V_P = \{p_j : j \in \mathbb{N}\}$ and edges $E_P = \{(p_j, p_{j+1}) : j \in \mathbb{N}\}$. The star-vertices are denoted $V_S = \{s_i : i \in \mathbb{N}\}$ and star edges $E_S = \{(p_j, s_{\alpha_j}) : j \in \mathbb{N}\}$.

Definition 4.10. We define an *infinite path-star graph* $R^\alpha = (V, E)$ where $V = V_P \cup V_S$ and $E = E_P \cup E_S$ (see example in Figure 4.2). We define the corresponding *path-star class* $\mathcal{R}^\alpha$ to be the finite induced subgraphs of R^α.

Any graph $G \in \mathcal{R}^\alpha$ can be witnessed by an embedding $\phi(G)$ into the infinite graph R^α. To simplify the presentation we will associate G with a particular embedding in R^α depending on the context.

To avoid confusion when referring to different types of path, we will refer to the *class-path* when referring to the (infinite) path of the path-star class, or a *path component* when referring to a finite section of it. A path component induced by the vertices $\{p_j, p_{j+1}, \ldots, p_{j+k}\}$ we denote $I_{[j,j+k]}$. We use the shorthand m*-path-vertex* for a vertex in the class-path corresponding to the letter m in α.

In addition, if α is a nested word over alphabet $\mathcal{A}$, and $\alpha^{\mathcal{S}}$ a nested subword restricted to the sub-alphabet $\mathcal{S}$, then we refer to *base star-vertices* and *base path-vertices* for vertices that correspond to base letters in $\mathcal{S}$ and *branch star-vertices* and *branch path-vertices* for vertices that correspond to branch letters in $\mathcal{S}$. These vertices will depend on the choice of $\mathcal{S}$.

4.3.1 Path-star classes with unbounded tree-width and clique-width

Throughout this section let $\mathcal{A}$ be an alphabet and α be an infinite word over $\mathcal{A}$. Let $\mathcal{A}^\alpha \subseteq \mathcal{A}$ be the set of letters in $\mathcal{A}$ that appear an infinite number of times in α. That is, these are the letters of $\mathcal{A}$ corresponding to the infinite stars in R^α.

Recalling the definition of t-sails (Definition 1.30), we have:

Lemma 4.11. *If G is a t-sail for positive integer $t \geqslant 2$ then $tw(G) \geqslant t-1$.*

Proof. Contracting star-vertex s_i with the vertices of path P_i for $1 \leqslant i \leqslant t$ gives a K_t-minor, then using Lemma 1.2 and the fact that $tw(K_t) = t-1$ we have $tw(G) \geqslant t-1$. □

It follows that hereditary graph classes containing arbitrarily large t-sails have unbounded tree-width.

Theorem 4.12. *If $\mathcal{A}^\alpha$ is infinite then the graph class $\mathcal{R}^\alpha$ has unbounded tree-width and clique-width.*

Proof. We will show that $\mathcal{R}^\alpha$ contains a t-sail for all t and thus has unbounded tree-width by Lemma 4.11. As $\mathcal{R}^\alpha$ has arboricity two it follows from Theorem 1.16 that $\mathcal{R}^\alpha$ also has unbounded clique-width.

Let $\mathcal{A}^\alpha = \{i_1, i_2, \dots\}$. For any $t \in \mathbb{N}$ we can create a set of t factors of α as follows.

Let $I_1 = \{j\}$, where j is the position of the first occurrence of letter i_1 in α. For $2 \leqslant k \leqslant t$ let $I_k = [x, y]$ be the next interval beyond I_{k-1} where α_{I_k} contains all of $i_1, \dots, i_k$. Such intervals can always be found because the letters in $\mathcal{A}^\alpha$ repeat infinitely in α. This gives us a set of t disjoint factors of α, $\{\alpha_{I_k} : 1 \leqslant k \leqslant t\}$.

Defining the vertex set $V^t = \{p_i : i \in \cup_{1 \leqslant k \leqslant t} I_k\} \cup \{s_{i_k} : 1 \leqslant k \leqslant t\}$ means $R^\alpha[V^t]$ is a t-sail and the result follows. □

4.3.2 Path-star classes are not subclasses of circle graphs

We showed in Section 1.3.3 that circle graphs play an important role in the study of rank-width and vertex-minor closed graph classes. More recently, in [39], the authors describe the unavoidable induced subgraphs of circle graphs with large tree-width. To distinguish the results in this Chapter from those in [39] we show that path-star graph classes are not subclasses of circle graphs.

Let α be a word over an alphabet with at least two letters. We will call a factor of α that starts with one letter i and ends with another j, with no other occurrences of either letter in the factor, an *(i, j)-alternance*. If G is a graph in the path-star class $\mathcal{R}^{\alpha^{\{1,2\}}}$ induced by the two stars s_1 and s_2 and a path component, we show that it is not possible to construct a chord representation of G when the sequence in α corresponding to the

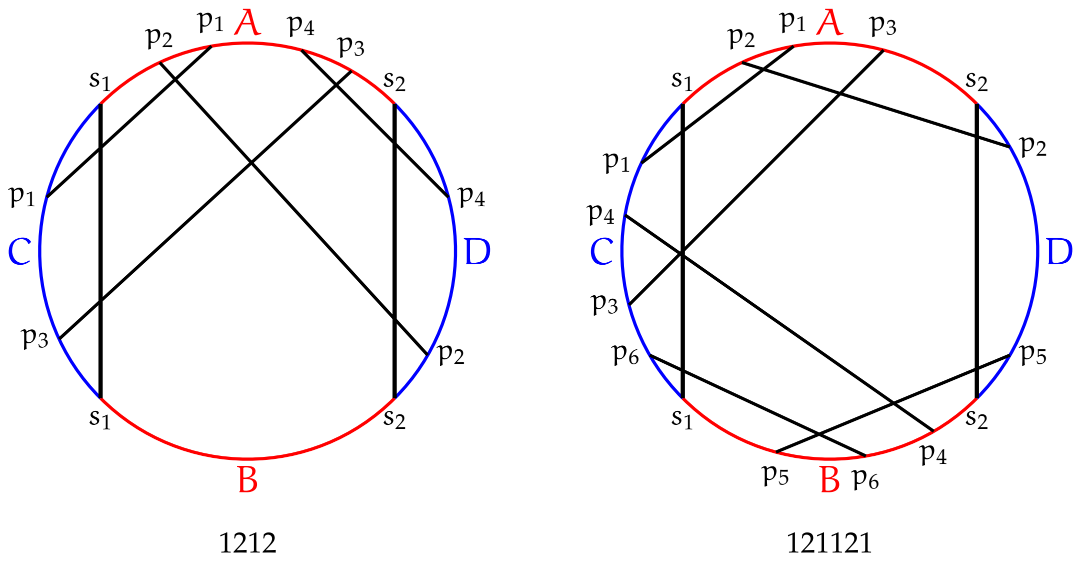


Figure 4.3: Chord representations of circle graphs 1212 and 121121

path component has more than four $(1,2)$-alternances, i.e., G is not a circle graph. For example, the word 11212221112 alternates 5 times between 1 and 2 and therefore does not represent a circle graph.

We may refer to G by name or by α letter sequence (e.g. $G = 1221221$). We will always label the path vertices of G starting with p_1 so that 1221221 has path vertices $p_1, \dots, p_7$.

Theorem 4.13. *If there are two letters in the word α that alternate more than four times then the graph class $\mathcal{R}^\alpha$ is not a subclass of circle graphs.*

Proof. Every graph in $\mathcal{R}^{\alpha^{\{1,2\}}}$ is a vertex-minor of a graph in $\mathcal{R}^\alpha$. As circle graphs are vertex-minor closed, if $\mathcal{R}^\alpha$ is a subclass of circle graphs then so is $\mathcal{R}^{\alpha^{\{1,2\}}}$. Thus, if we can find a graph in $\mathcal{R}^{\alpha^{\{1,2\}}}$ that is not a circle graph then we are done. Also note that there is a $1-1$ correspondence between the $(1,2)$-alternances in α and $\alpha^{\{1,2\}}$.

Suppose, for a contradiction, that there exists a factor β of $\alpha^{\{1,2\}}$ in which the letters 1 and 2 alternate more than four times, with the property that the graph G induced by the stars s_1 and s_2 and the path component corresponding to the factor β is a circle graph.

We try to construct a chord representation for G – see examples in Figure 4.3. Without loss of generality, we assume that the first letter of β is 1 and the second 2.

Note that the chords representing s_1 and s_2 do not cross, shown as vertical lines in Figure 4.3. Designating the arcs between s_1 and s_2 A and B (shown in red), and the arcs bounded by s_1 and s_2 C and D respectively (shown in blue), note that every path vertex adjacent to s_1 must be represented by a chord with one end in C and the other in either A or B, and similarly for s_2, a chord with one end in D and the other in either A or B. Therefore, if p_1 and p_2 are the two chords representing the first $(1,2)$-alternance then they must cross and both have an end in either A or B. Without loss of generality, let them both have an end in A. We will show that there cannot be many consecutive alternances happening in the same sector.

Notice that for any i, the chords representing $p_{i+2}, p_{i+3}, \ldots$ must all be on the same side of the chord representing p_i as none of them can cross this chord. Also notice that if $i < j < k$ and $\alpha_{p_i} = \alpha_{p_j} = 1$, $\alpha_{p_k} = 2$ then the chord for p_j must be situated on the s_2 side of the chord for p_i to accommodate the next alternance (and likewise with 1 and 2 reversed).

Suppose that no path-vertex chord has an end in B. If p_3 is a 1 (i.e., a second $(1,2)$-alternance) then its chord must cross only s_1 and p_2. If this is on the 'non-s_2' side of p_1 then this prevents any further alternance since the path is blocked from star s_2 by chord p_1. So for there to be a third alternance, p_3 must be on the s_2 side as shown in the 1212 example in Figure 4.3.

The chord representing path-vertex p_4 cannot cross p_1 or p_2. Furthermore, if it is on the 'non-s_1' side of p_2 then this prevents any further alternance since the path is blocked from star s_1 by chord p_2. Hence, without using arc B, we can have at most three $(1,2)$-alternances.

Now suppose that a path-vertex chord may have an end in B. We may have at most two alternances through A before switching to B, as if we start with three alternances through A, as shown in the 1212 example in Figure 4.3, then p_4 is blocked from B.

If we switch to B after two alternances then p_4 is a 1 with chord ends in C and B. It is possible to have at most two alternances through B before we reach p_6 which is blocked by p_4, as shown in the 121121 example in Figure 4.3, and thereafter no further alternance is possible either via A or B.

It follows that the maximum number of alternances possible is four. By assumption, the letters in β alternate more than four times and hence we cannot construct a chord representation of G, and we have a contradiction. □

4.4 Nested path-star hereditary graph classes

We now focus on path-star graph classes defined by nested words.

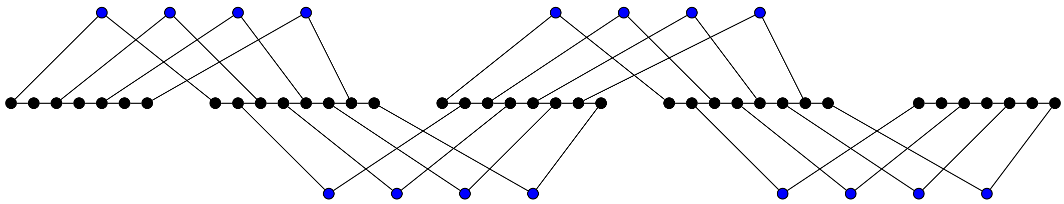


Figure 4.4: Example of a subdivision of a 4×4-wall embedded in a path-star graph (star-vertices blue)

4.4.1 Nested path-star classes are KKW-free

It is quite possible for path-star graph classes to contain a large wall – see an example in Figure 4.4. However, we show that path-star graph classes created from nested words (see Definition 4.2) are KKW-free.

Observe that if α is a nested word then any connected graph G in $\mathcal{R}^\alpha$ that does not contain a base-path-vertex contains only star- and path-vertices corresponding to a single branch of α, since it is not possible to have a path connecting vertices corresponding to two different branches that does not contain a base-path-vertex.

Lemma 4.14. *If α is a nested word and G is a graph in $\mathcal{R}^\alpha$ that does not contain a base-path-vertex then any induced hole in G must contain exactly two star-vertices.*

Proof. Clearly, a hole must contain at least one star-vertex, otherwise it would be a path. Suppose it contained only one star-vertex, say s_x, then α would have a factor $x \dots x$ containing no base letter since G has no base-path-vertices. Since the vertices only correspond to a single branch of α and branch letters appear in branch-order, such a factor does not exist so we have a contradiction.

Suppose our hole contains three or more star-vertices, say s_x, s_y and s_z where x, y and z appear in this order in a branch. This requires the three stars to be connected by path segments corresponding to branch sequences $x \dots y$, $x \dots z$ and $y \dots z$ in α. As G is single-branch then the sequence $x \dots z$ must contain the letter y. The corresponding path-vertex must be adjacent to s_y creating a chord in the cycle, so it is not a hole. A contradiction.

Therefore, the only possibility is that any hole in G must contain exactly two star-vertices. □

We will call a graph consisting of five 'bricks' of a wall, as shown with numbered vertices in Figure 4.5, a 5-*wall*.

Notice that a 5-wall contains exactly eleven induced chordless cycles or holes – which we will call $H_1, \dots, H_{11}$ (shown in Figure 4.5) with vertex sets $V_1, \dots, V_{11}$ respectively.

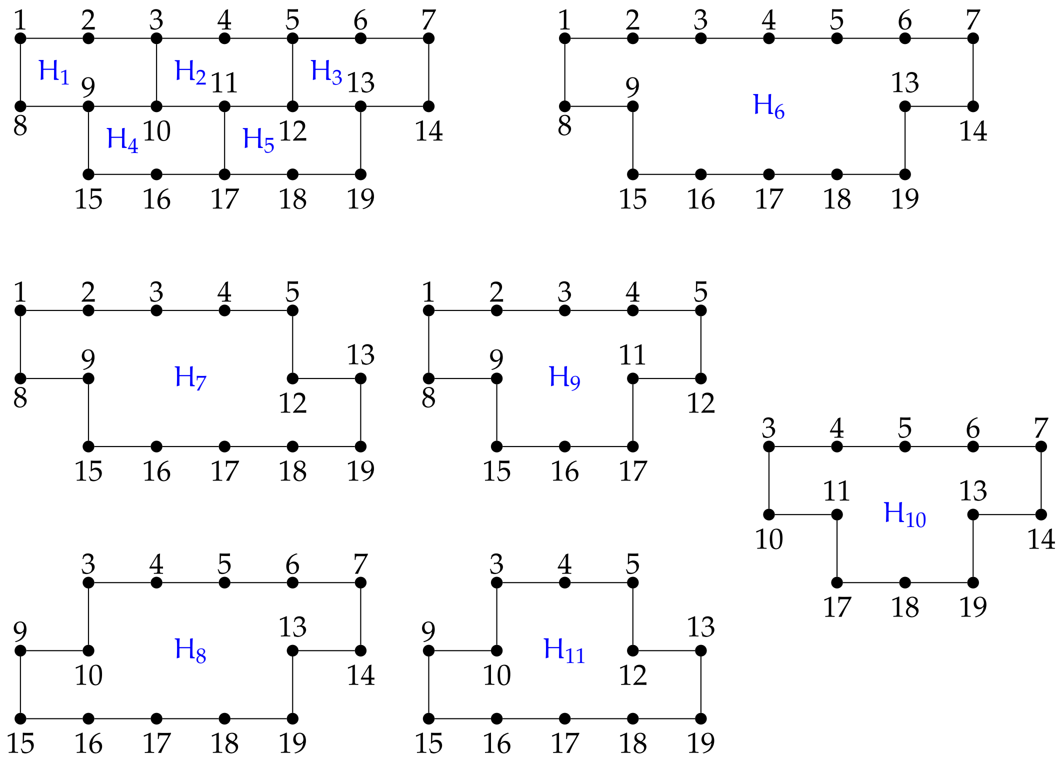


Figure 4.5: A 5-wall (top left) together with its eleven holes $H_1, \dots, H_{11}$

Lemma 4.15. *If α is a nested word then every subdivision of a 5-wall in $\mathcal{R}^\alpha$ contains a base-path-vertex.*

Proof. Suppose, for a contradiction, that $\mathcal{R}^\alpha$ contains a graph G that is a subdivision of a 5-wall that does not contain a base-path vertex. From Lemma 4.14 there must be precisely two star-vertices in each vertex-set $V_1, \dots, V_{11}$ in G.

Firstly, we consider V_1 and V_3. Notice that if there are two star-vertices in $V_1 \cap V_8 = V_1 \cap V_{11}$ then there can be no star-vertices in $V_3 \subset (V_8 \cup V_{11})$, a contradiction. So there must be at least one star vertex in $V_1 \setminus V_{11} \subset V_6$ (set W_1) and by symmetry at least one star-vertex in $V_3 \setminus V_{11} \subset V_6$ (set W_2). Since these are disjoint subsets of V_6 there are no other star-vertices in V_6.

Observe that a degree three vertex in a path-star graph must either be a star-vertex or adjacent to a star-vertex. In particular, vertex 17 in Figure 4.5 is degree three, and combined with the fact that the vertices adjacent to it in V_6 are not star-vertices, as demonstrated above, means that there is a star-vertex in $(V_9 \cap V_{10}) \setminus V_6$ (set W_3, that is, the path segment 11 – 17).

One star-vertex is in $W_1 \subset V_9$, so another star-vertex is in $V_9 \setminus W_1$. Likewise, there is a star-vertex in $W_2 \subset V_{10}$, so another star-vertex is in $V_{10} \setminus W_2$. Both of these vertex sets include W_3, which from the previous paragraph must contain a star-vertex. Consequently, neither $V_9 \setminus (W_1 \cup W_3)$ nor $V_{10} \setminus (W_2 \cup W_3)$ contains a star-vertex.

Combining these sets, $(V_9 \cup V_{10}) \setminus (W_1 \cup W_2 \cup W_3)$ does not contain a star-vertex. But this contains all of V_2 except for vertex 11, so V_2 contains at most one star-vertex, a contradiction.

Therefore, it is not possible to construct G without a base-path vertex. □

Theorem 4.16. *If α is a nested word then $\mathcal{R}^\alpha$ is* KKW*-free.*

Proof. $\mathcal{R}^\alpha$ has arboricity two so does not contain K_5 or $K_{4,4}$.

We show that $\mathcal{R}^\alpha$ does not contain a subdivision of a $t \times t$ wall for t where α is $\mathfrak{b}$-nested and $\left\lfloor \frac{t}{10} \right\rfloor > \sqrt{\frac{\mathfrak{b}}{3}}$.

Suppose, for a contradiction, that $\mathcal{R}^\alpha$ contains a graph G that is a subdivision of a $t \times t$ wall for some t where $\left\lfloor \frac{t}{10} \right\rfloor > \sqrt{\frac{\mathfrak{b}}{3}}$. Fix some embedding of this wall into R^α, and let $\mathcal{S} \subseteq \mathcal{A}$ denote the letters whose star-vertices appear in this embedding.

Since α is $\mathfrak{b}$-nested, $\mathcal{S}$ has a base $\mathcal{B}$ where $|\mathcal{B}| \leqslant \mathfrak{b}$. For any $x \in \mathcal{S}$ there can be at most three x-path-vertices in G since s_x is a vertex of degree at most three. Hence, $V(G)$ can contain at most $3\mathfrak{b} < 9\left\lfloor \frac{t}{10} \right\rfloor^2$ base-path-vertices.

From Lemma 4.15 every induced subdivision of a 5-wall in G must contain a base-path-vertex. Using Figure 4.6 it is possible to pack at least nine vertex-disjoint 5-walls into a 10×10-wall, so our subdivision of a $t \times t$ wall must contain at least $9\left\lfloor \frac{t}{10} \right\rfloor^2$ disjoint subdivisions of a 5-wall.

Hence, allowing at least one base-path-vertex in each induced subdivision of a 5-wall, $V(G)$ must contain at least $9\left\lfloor \frac{t}{10} \right\rfloor^2$ base-path-vertices. But we know $V(G)$ contains at most $3\mathfrak{b} < 9\left\lfloor \frac{t}{10} \right\rfloor^2$ base-path-vertices, so we have a contradiction. Thus, $\mathcal{R}^\alpha$ cannot contain a subdivision of a $t \times t$ wall when $\left\lfloor \frac{t}{10} \right\rfloor > \sqrt{\frac{\mathfrak{b}}{3}}$.

A similar argument can be applied to a line graph of a subdivision of a $t \times t$ wall noting that each triangle in the line graph contains a star-vertex. □

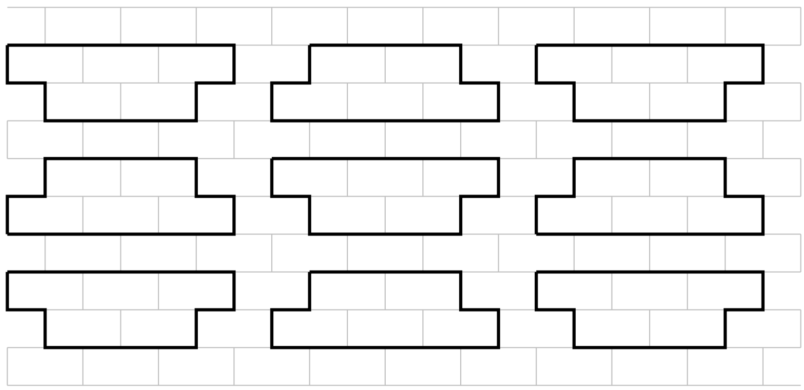


Figure 4.6: Nine 5-walls in a 10×10 wall

4.4.2 A nested path-star graph with large tree-width contains a large t-sail

A k-*block* in a graph G is a maximal set of at least k vertices no two of which can be separated in G by deleting fewer than k vertices. A k-block can be thought of as a highly connected part of a graph and has been used in a number of ways. In particular, Weißauer showed in [61] that for $k \geqslant 1$ every graph of tree-width at least $2k^2$ has a minor containing a k-block.

In [3] a more restricted type of k-block was introduced. A *strong* k-*block* in G is a set B of at least k vertices such that for every 2-subset $\{x, y\}$ of B, there exists a collection $\mathcal{P}_{x,y}$ of at least k distinct and pairwise internally disjoint paths in G from x to y, where for every two distinct 2-subsets $\{x, y\}, \{x', y'\} \subseteq B$ and every choice of paths $P \in \mathcal{P}_{x,y}$ and $P' \in \mathcal{P}_{x',y'}$ we have $P \cap P' = \{x, y\} \cap \{x', y'\}$.

We show that all t-sails for large t contain strong k-blocks for large k and that in nested path-star graph classes, strong k-blocks for large k only occur in graphs containing a t-sail for large t as an induced subgraph . We use this to conclude that a nested path-star graph has large tree-width if and only if it contains a t-sail for large t as an induced subgraph.

Lemma 4.17. *For any* $t \geqslant 1$ *a* t^3*-sail contains a strong* t*-block.*

Proof. Let $B = \{s_1, \ldots s_t\}$, i.e., the first t star-vertices. We claim B is a strong t-block.

For every 2-subset $\{s_x, s_y\}$ of B, we define the set $\mathcal{P}_{x,y}$ of t disjoint paths between s_x and s_y $(x < y)$ being the t class-path components numbered $\{(x-1)t^2 + (y-1)t, (x-1)t^2 + (y-1)t + 1, \ldots, (x-1)t^2 + (y-1)t + t - 1\}$.

For every two distinct 2-subsets $\{x, y\}, \{x', y'\} \subseteq B$ and every choice of paths $P \in \mathcal{P}_{x,y}$ and $P' \in \mathcal{P}_{x',y'}$ we have $P \cap P' = \{s_x, s_y\} \cap \{s'_x, s'_y\}$. Hence B is a strong t-block. □

To show that in nested path-star graph classes a large strong k-block contains a large t-sail as an induced subgraph, we explore the structure of nested words further.

Let $\mathcal{S}_1$ be the infinite set of all letters appearing in the $\mathfrak{b}$-nested word α. We write $\mathcal{B}_1$ for the base of α. Define nested subwords $\alpha^{\mathcal{S}_i}$ for $i \geqslant 2$ by $\mathcal{S}_i = \mathcal{S}_{i-1} \setminus \mathcal{B}_{i-1}$ where $\mathcal{B}_i$ is the base of $\alpha^{\mathcal{S}_i}$. Note that each $\mathcal{B}_i$ contains at least one and at most $\mathfrak{b}$ letters.

For any positive integer t, define a $\mathcal{B}^t$*-factor* as a sequence of α containing only letters from $\cup_{i=1}^t \mathcal{B}_i$, and at least one letter from each $\mathcal{B}_i$ for $1 \leqslant i \leqslant t$. Likewise, define a $\mathcal{S}^t$*-factor* as a sequence of α containing only letters from a single branch of $\alpha^{\mathcal{S}_t}$.

Observation 4.18. *For any positive integer* t*, a nested word* α *alternates between* $\mathcal{B}^t$*-factors and* $\mathcal{S}^t$*-factors.*

Proof. In a branched word (see Definition 4.1), one or more base letters must appear immediately before and immediately after a branch, so the statement is true for $t = 1$. Using induction, assume the statement is true for $t = n-1$ so that α alternates between $\mathcal{B}^{n-1}$-factors and $\mathcal{S}^{n-1}$-factors.

Suppose, in α, we have a sequence uvw where u and w are $\mathcal{S}^{n-1}$-factors and v is a $\mathcal{B}^{n-1}$-factor. From Definition 4.2, in $\alpha^{\mathcal{S}_n}$ each of u and w are reduced to a factor that contains at most one $\mathcal{S}^n$-factor (since from bullet point three of the definition it is not possible to get two branches in $\alpha^{\mathcal{S}_n}$ out of one branch in $\alpha^{\mathcal{S}_{n-1}}$) and v completely disappears.

As $\alpha^{\mathcal{S}_n}$ is branched we cannot have two adjacent $\mathcal{S}^n$-factors so there must be a letter(s) from $\mathcal{B}_n$ between them. Combining this letter(s) with the $\mathcal{B}^{n-1}$-factor v, gives us a $\mathcal{B}^n$-factor between every pair of $\mathcal{S}^n$-factors, so the statement is true for $t = n$.

Therefore, from the induction hypothesis, the observation follows. □

Observation 4.19. *Let* $\mathcal{H} = \{h_1, h_2, \dots\}$ *be a branch set of* $\alpha^{\mathcal{S}_t}$ *where* $h_1 < h_2 < \dots$*. Then for any* $x < y$*, between any occurrence of* h_x *and* h_y *in* α *there is either a factor* $h_1 \dots h_{x-1}$ *or* $h_{x+1} \dots h_{y-1}$*.*

Proof. This follows from the fact that branches must start with the first letter in the branch set and must appear in branch order. □

Lemma 4.20. *Let* G *be a graph from a path-star class defined by a* $\mathfrak{b}$*-nested word* α *over the infinite alphabet* $\mathcal{A}$*. If* G *contains a strong* k*-block, where* $k \geqslant \max\{t\mathfrak{b}^t + t\mathfrak{b}, t(\mathfrak{b}+2)+2\}$ *for some integer* $t \geqslant 1$*, then it also contains a* t*-sail as an induced subgraph.*

Proof. Fix some embedding of G into R^α, and let $\mathcal{S}_1 \subseteq \mathcal{A}$ denote the letters whose star-vertices appear in this embedding. For any two vertices in a strong k-block there must be k internally disjoint paths between them. The vertices in this strong k-block must be star-vertices since only star-vertices can have degree greater than three, so let $\mathcal{L} \subseteq \mathcal{S}_1$ be the letters corresponding to the vertices of the strong k-block in G where $k \geqslant t\mathfrak{b}^t + t\mathfrak{b}$.

Let subword $\alpha^{\mathcal{S}_1}$ have base $\mathcal{B}_1$, and define subwords $\alpha^{\mathcal{S}_i}$ for $i \geqslant 2$ by $\mathcal{S}_i = \mathcal{S}_{i-1} \setminus \mathcal{B}_{i-1}$ with base $\mathcal{B}_i$.

Observe that each star-vertex corresponding to a letter in $\mathcal{S}_1 \setminus \mathcal{L}$ appears in at most one of the internally disjoint paths between two vertices of the strong k-block as otherwise the paths would not be disjoint.

As $k \geqslant t(\mathfrak{b}+2)+2$, and there are at most $t\mathfrak{b}$ letters in $\cup_{i=1}^{t}\mathcal{B}_i$, $\mathcal{L}$ contains at least $(2t+2)$ letters in $\mathcal{S}_{t+1}$. Therefore, either there exists a pair $x, y \in \mathcal{L}$ that are in different branch sets of $\alpha^{\mathcal{S}_t}$ (Case 1) or there are at least $(2t+2)$ letters in the same branch set of $\alpha^{\mathcal{S}_t}$ (Case 2).

Case 1: (There exists a pair $x, y \in \mathcal{L}$ in different branch sets of $\alpha^{\mathcal{S}_t}$.) From Observation 4.18, between every occurrence of x and occurrence of y in α there is a $\mathcal{B}^t$ factor.

At most one of the disjoint paths from s_x to s_y can pass through each star-vertex associated with a letter in $\cup_{i=1}^{t}\mathcal{B}_i$ (i.e., at most $t\mathfrak{b}$ paths). This leaves $k - t\mathfrak{b}$ paths that do not pass through such a star-vertex. The remaining disjoint paths must all include a set of consecutive class-path-vertices corresponding to a $\mathcal{B}^t$ factor in α (Observation 4.18).

Given a collection of $\mathcal{B}^t$ factors of size $k - t\mathfrak{b}$, using the pigeonhole principle, as $k \geqslant t\mathfrak{b}^t + t\mathfrak{b}$, there are at least $\frac{k-t\mathfrak{b}}{\mathfrak{b}^t} \geqslant t$ of them that contain the same letter from each set $\mathcal{B}_i$, $1 \leqslant i \leqslant t$. Call this set of at least t letters $\mathcal{T} \subset \cup_{i=1}^{t}\mathcal{B}_i$. It follows that at least t of the disjoint paths from s_x to s_y contain a component of the class-path incorporating a path-vertex corresponding to each letter in $\mathcal{T}$ – let us call these components $I_1, \dots I_t$.

Case 2: (There exists at least $(2t+2)$ letters in the same branch set of $\alpha^{\mathcal{S}_t}$.) Let the $(2t+2)$ letters come from branch $\mathcal{H} = \{h_1, h_2, \dots\}$ where $h_1 < h_2 < \dots$. Since there are at least $(2t+2)$ letters, we can choose the $(t+1)$-th and $(2t+2)$-th letters as h_x and h_y so that $t < x < x + t < y$. Using Observation 4.19, between every occurrence of h_x and occurrence of h_y in α there is a set of t consecutive letters from $\mathcal{H}$ (either the first t letters in $\mathcal{H}$ or the (at least) t letters between h_x and h_y in the $\mathcal{H}$ order).

Denoting s_x and s_y as the two star-vertices in the k-block corresponding to letters h_x and h_y, at most one of the disjoint paths from s_x to s_y can pass through each star-vertex associated with one of these $2t$ letters in $\mathcal{H}$. This leaves $k - 2t$ disjoint paths that do not pass through such a star-vertex. The remaining disjoint paths must all include a set of consecutive class-path-vertices corresponding to one of our two sets of consecutive letters from $\mathcal{H}$ (Observation 4.19).

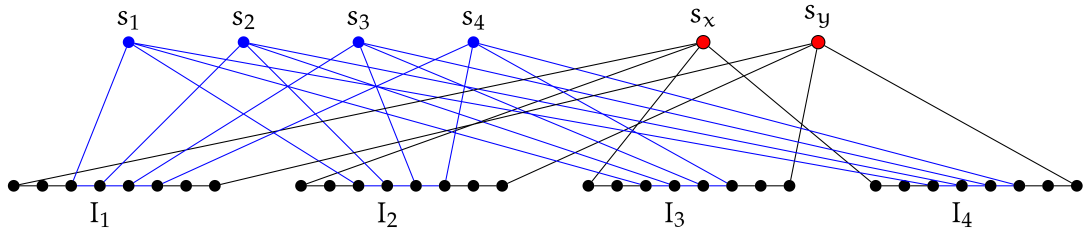


Figure 4.7: A 4-sail (blue) with s_x and s_y two nodes in a k-block

From the pigeonhole principle at least $\frac{k-2t}{2} \geqslant t$ of the disjoint paths must correspond to the same set of t consecutive letters from $\mathcal{H}$. Call this set of letters $\mathcal{T} \subseteq \mathcal{H}$. It follows that at least t of the disjoint paths from s_x to s_y contain a component of the class-path incorporating a path-vertex corresponding to each letter in $\mathcal{T}$ – let us call these components $I_1, \dots I_t$.

In either Case 1 or Case 2, we have path components in G, $I_1, \dots I_t$, each containing path-vertices corresponding to each letter in $\mathcal{T}$. Given that $\mathcal{T} \subset \mathcal{S}_1$ the stars $S_\mathcal{T}$ corresponding to letters in $\mathcal{T}$ are all in V[G]. Therefore, the graph $G\left[\bigcup_{i=1}^{t} V[I_i] \cup S_\mathcal{T}\right]$ contains a forest of t paths and a forest of t stars sufficient to fulfil the definition that it contains a t-sail as an induced subgraph [see example in Figure 4.7]. □

Letting $\mathbb{B}_k$ be the class of all graphs with no strong k-block and remembering that the k-*basic obstructions* are (1) the complete graph K_k, (2) the complete bipartite graph $K_{k,k}$, (3) a subdivision of the $(k \times k)$-wall and (4) a line graph of a subdivision of the $(k \times k)$-wall, we use the following result:

Theorem 4.21 ([3]). *For every integer $k \geqslant 1$ there exists a positive integer $w(k)$ such that every graph in $\mathbb{B}_k$ with tree-width more than $w(k)$ contains an induced subgraph isomorphic to one of the k-basic obstructions.*

Theorem 4.22. *If α is a nested word then for every $t \geqslant 1$ there is a positive integer valued function $f_\alpha(t)$ such that every graph in $\mathcal{R}^\alpha$ of tree-width at least $f_\alpha(t)$ contains a t-sail as an induced subgraph.*

Proof. Let $k = \max\{t\mathfrak{b}^t + t\mathfrak{b}, t(\mathfrak{b}+2)+2\}$ and $f_\alpha(t) = w(k)$ as defined by Theorem 4.21. Suppose for graph $G \in \mathcal{R}^\alpha$, we have $tw(G) \geqslant f_\alpha(t)$. Then by Theorem 4.21, G cannot be in $\mathbb{B}_k$ because by Theorem 4.16 $\mathcal{R}^\alpha$ is KKW-free, G does not contain a k-basic obstruction, and therefore, G contains a strong k-block. It follows by Lemma 4.20 that G contains an induced subgraph isomorphic to a t-sail. □

4.4.3 Nested path-star classes are infinitely defined

As previously mentioned, the list of minimal forbidden induced subgraphs in a hereditary class may be finite or infinite. Theorem 1.18 characterises tree-width in finitely defined graph classes. We do not have an equivalent characterisation of tree-width for infinitely defined classes. However, we can show that one consequence of Theorem 1.18 is the following:

Theorem 4.23. *A hereditary class of graphs of unbounded tree-width that is* KKW*-free is infinitely defined.*

Proof. Let $\mathcal{C}$ be a hereditary class of graphs of unbounded tree-width that is KKW-free, so that it excludes a subdivision of a $t \times t$-wall and line graph of a subdivision of a $t \times t$-wall for some $t \in \mathbb{N}$.

For a contradiction suppose $\mathcal{C}$ is finitely defined with minimal forbidden induced subgraphs $\mathcal{F} = \{F_1, \dots, F_n\}$ for some $n \in \mathbb{N}$. As $\mathcal{C}$ has unbounded tree-width, then from Theorem 1.18 $\mathcal{F}$ either does not contain a tripod or does not contain the line graph of a tripod. Suppose it does not contain a tripod (i.e. all tripods are in $\mathcal{C}$).

Let d denote the maximum distance between any two vertices of degree three or more for any tree (other than a tripod) in $\mathcal{F}$, and let m denote the maximum girth of any induced cycle in any graph in $\mathcal{F}$.

Let W be a subdivision of a $t \times t$-wall with more than $\max(d, m)$ degree two vertices on the paths between degree three vertices. This contains none of the minimal forbidden induced subgraphs in $\mathcal{F}$ since every induced cycle of W has length more than m and every induced tree of W has distance greater than d between degree three vertices, hence, $W \in \mathcal{C}$. A contradiction of the assumption that $\mathcal{C}$ excludes a subdivision of a $t \times t$-wall.

A similar argument applies if $\mathcal{F}$ does not contain the line graph of a tripod, thus $\mathcal{C}$ must be infinitely defined. □

Corollary 4.24. *A path-star class defined by a nested word is infinitely defined.*

Proof. From Theorem 4.16 all path-star classes defined by a nested word are KKW-free so from Theorem 4.23 are infinitely defined. □

4.5 Minimal path-star classes of unbounded tree-width

As discussed in Section 1.7, a hereditary class of graphs of unbounded clique-width must contain a minimal class or an infinite antichain of unbounded clique-width (or both). In this section, we show that a hereditary class of graphs of unbounded clique-width (a) of bounded vertex degree, or (b) that has an excluded minor or (c) is a path-star class defined by a nested word, does not contain a minimal class. Therefore, by default, these classes contain an infinite antichain of unbounded clique-width. We revisit such objects in Chapter 6.

4.5.1 Hereditary graph classes of bounded vertex degree or with an excluded minor do not contain a minimal subclass

The structure of a wall allows us to delete vertices and leave the fundamental structure intact, ignoring subdivisions, and this quality is used in the following:

Lemma 4.25. *A subdivision (or line graph of a subdivision) of a $W_{kt\times kt}$ wall for $k, t \geqslant 1$ contains an induced subgraph isomorphic to a subdivision (or line graph of a subdivision) of a $W_{t\times t}$ wall that does not contain a cycle smaller than C_{8k-2} (other than C_3 in the case of the line graph).*

Proof. Let $G = (V, E)$ be a subdivision of a $W_{kt\times kt}$ wall. Let $V^3 \subseteq V$ be the set of degree three vertices in G, together with the equivalent degree two vertices from the perimeter 'bricks' (or holes) that would be degree three if the wall was extended, so that every brick in G contains six vertices in V^3. An induced subgraph G' isomorphic to a subdivision of a $W_{t\times t}$ wall can be constructed by overlaying a lattice of $k \times k$ sub-walls, the new 'bricks', and deleting all vertices of G internal to every new brick, as shown in the example in Figure 4.8 where $k = 2$ and $t = 4$. Each new brick contains $8k - 2$ vertices from V^3, and thus G' does not contain a cycle smaller than C_{8k-2}.

An identical argument works if G is the line graph of a subdivision of a $W_{kt\times kt}$ wall. □

Theorem 4.26. *If $\mathcal{C}$ is a hereditary class of graphs of bounded vertex degree or that has an excluded minor then it does not contain a minimal class.*

Proof. If $\mathcal{D}$ is a minimal hereditary subclass of $\mathcal{C}$ then by Theorems 1.26 or 1.27, as it has unbounded tree-width, it contains (as a member of the class) a graph G which is isomorphic to a subdivision (or line graph of a subdivision) of a $W_{kt\times kt}$ wall for arbitrarily large k and t.

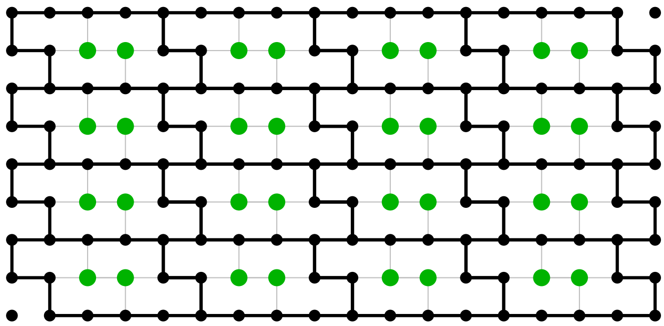


Figure 4.8: An 8×8 wall containing a subdivision of a 4×4 wall after large (green) vertex deletion

Suppose C_m $(m > 3)$ is the shortest cycle in G. Set $k > \frac{m+2}{8}$. Then from Lemma 4.25 G contains as an induced subgraph a subdivision (or line graph of a subdivision) of a $W_{t \times t}$ wall that does not contain a cycle smaller than C_{8k-2} which is longer than C_m (other than C_3).

But now the proper hereditary subclass $\mathcal{D} \cap \mathtt{Free}(C_m)$ contains a subdivision of $W_{t \times t}$ for arbitrarily large t, so $\mathcal{D} \cap \mathtt{Free}(C_m)$ also has unbounded tree-width, which contradicts $\mathcal{D}$ being minimal. □

4.5.2 Nested path-star hereditary graph classes do not contain a minimal subclass

We show that no path-star hereditary graph class defined by a nested word contains a minimal subclass.

Lemma 4.27. *Let α be a $\mathfrak{b}$-nested word over an infinite alphabet $\mathcal{A}$. Then for integers $t \geqslant 2$, $m > 3$ and $T \geqslant 2t + m\mathfrak{b}$, a T-sail in $\mathcal{R}^\alpha$ with smallest cycle C_m contains an induced subgraph isomorphic to a t-sail which does not contain C_m.*

Proof. Let G be a T-sail in $\mathcal{R}^\alpha$. Fix some embedding of G into R^α, and let $\mathcal{S}_1 \subseteq \mathcal{A}$ denote the letters whose star-vertices appear in this embedding.

Let subword $\alpha^{\mathcal{S}_1}$ have base $\mathcal{B}_1$, and define nested subwords $\alpha^{\mathcal{S}_i}$ for $2 \leqslant i \leqslant m$ by $\mathcal{S}_i = \mathcal{S}_{i-1} \setminus \mathcal{B}_{i-1}$ with base $\mathcal{B}_i$ (Note $|\mathcal{B}_i| \leqslant \mathfrak{b}$ for all $i = 1, 2, \ldots$).

Let G' be the subgraph of G induced by the vertices of G excluding the star-vertices corresponding to the letters of $\cup_{i=1}^{m} \mathcal{B}_i$ (at most $m\mathfrak{b}$) and excluding the star-vertices corresponding to alternate letters in the branch sets of $\mathcal{S}_m$ (i.e., half the remaining star-vertices).

We claim G' contains an induced subgraph isomorphic to a t-sail which does not contain C_m.

That G' contains an induced subgraph isomorphic to a t-sail follows from the fact that it contains at least $\frac{T-mb}{2} \geqslant t$ of the star-vertices of G and all the path-vertices from the path components of G.

Suppose that G' contains an m-cycle with only one star-vertex, say s_x. Then the path-vertices adjacent to s_x in the cycle must both correspond to a branch letter x of $\mathcal{S}_m$. The rest of the cycle must consist of path-vertices corresponding to a factor $x \ldots x$ of α. Using Observations 4.18 and 4.19 there must be at least one base path-vertex from each of the m base sets $\mathcal{B}_i$ in the cycle so it has more than m vertices, a contradiction.

Suppose that G' contains an m-cycle with three (or more) star-vertices, s_x, s_y and s_z. This must contain path components corresponding to α factors of the form $x \ldots y$, $x \ldots z$ and $y \ldots z$ (or their reverse). These factors must be contained in branches of $\alpha^{\mathcal{S}_m}$, and hence x, y and z must be from the same branch set, otherwise the cycle would have more than m vertices for the same reason as for the one star-vertex case. But now, assuming without loss of generality that their branch order is $x < y < z$, then there must be a y in the $x \ldots z$ factor and a shorter cycle exists, which contradicts the fact that the shortest cycle in G is of length m.

Lastly, suppose that G' contains an m-cycle with precisely two star-vertices, s_x and s_z. This must contain path components corresponding to two α factors of the form $x \ldots z$ (or the reverse). Hence, x and z are from the same branch set. In fact they must be consecutive letters from the same branch set, since if there was another letter y between them in the branch order then there would be a shorter cycle than C_m, either containing the two stars s_x and s_y or the two stars s_y and s_z. But the construction of G' requires the removal of star-vertices corresponding to alternate letters in the branch sets of $\mathcal{S}_m$, so x and z cannot be consecutive branch letters, a contradiction. Hence G' does not contain C_m. □

Theorem 4.28. *If $\mathcal{R}^\alpha$ is a path-star hereditary class of graphs defined by a nested word α then it does not contain a minimal class.*

Proof. If $\mathcal{D}$ is a minimal subclass of $\mathcal{R}^\alpha$ then by Theorem 4.22 for every positive integer T, $\mathcal{D}$ contains a T-sail.

Suppose the shortest cycle in $\mathcal{D}$ is C_m ($m > 3$). Then from Lemma 4.27 for any positive integer t there exists a positive integer T such that any T-sail in $\mathcal{D}$ contains an induced subgraph isomorphic to a t-sail which does not contain a C_m cycle. Thus the subclass

$\mathcal{D} \cap \mathsf{Free}(C_m)$ still contains a t-sail for arbitrarily large t and has unbounded tree-width, which contradicts $\mathcal{D}$ being minimal. □

4.5.3 Further thoughts

This study is, as far as we know, the first time an attempt has been made to use combinatorics on words in the study of treewidth. We believe the results are sufficient enough to justify further use of this technique. Likewise, path-star hereditary graph classes seem to be significant in respect of the study of tree-width and clique-width in sparse graph classes and warrant a more thorough study.

We have shown that path-star graph classes defined by nested words exclude large walls and large line graphs of walls. However, we have not resolved whether there are other such words, or whether, if we forbid a large wall and a large line graph of a wall in a path-star graph class $\mathcal{R}^\alpha$, then α contains a large nested subword.

Chapter 5

Dense Classes : An Alternative Framework

5.1 Introduction

In Chapter 2 we constructed a framework for minimal hereditary classes of unbounded clique-width that incorporated most of the known minimal classes. One exception was a class of ’power graphs’ introduced by Lozin, Razgon and Zamaraev [49] and proved to be a minimal class by Dawar and Sankaran [29]. Taking this example, we construct an alternative framework for minimal hereditary classes of unbounded clique-width that includes the power graphs (but not the minimal classes in the framework from Chapter 2).

New ’path-clique’ graph classes are created that consist of the induced subgraphs of an infinite graph whose vertex set is indexed by the natural numbers. These vertices are partitioned into an infinite number of sets, each set being infinite in number. Edges consist of the symmetric difference between the edge sets of two graphs defined on the vertices:

1. an infinite path in vertex index order, and
2. a ’biclique graph’ where each set of the partition is a clique or independent set and such sets are complete or anticomplete to each other.

Using local complementation we create a process to sparsify a dense graph, which helps us identify a new family of hereditary classes of unbounded clique-width, an uncountable number of which are minimal. We draw an interesting parallel between these dense path-clique minimal classes of unbounded clique-width and sparse path-star classes (see Chapter 4) which have unbounded clique-width but do not contain a minimal class.

5.2 Constructing the framework for path-clique graph classes

5.2.1 Path-clique graph classes

In the following, the class-path P is identical to one defined for path-star classes in Section 4.3. However, other definitions are different to Chapter 4 since we are dealing with cliques and not stars.

We use a pair of objects (α, β) to define our graph classes, where

(a) α is an infinite word from the alphabet $\mathcal{A} = \mathbb{N}$ where each letter occurs an infinite number of times, and

(b) β is a symmetric set of pairs of natural numbers (i, j).

We first define two auxiliary graphs:

The infinite *class-path* $P = (V_P, E_P)$ has vertices $V_P = \{p_j : j \in \mathbb{N}\}$ and edges $E_P = \{p_j p_{j+1} : j \in \mathbb{N}\}$. A consecutive set of vertices $\{p_j, p_{j+1}, \ldots, p_{j+k}\}$ we call an *interval* denoted $I_{[j,j+k]}$.

The *biclique graph* $B^{(\alpha,\beta)} = (V_P, E_B)$, defined on the same vertices, has edges $E_B = \{p_i p_j : i < j, (\alpha_i, \alpha_j) \in \beta\}$

We define a partition of V_P with vertex sets or 'bags' $Q_j = \{p_i : \alpha_i = j\}$. Thus, for bag Q_i, $B^{(\alpha,\beta)}[Q_i]$ is a clique if $(i,i) \in \beta$ and an independent set if $(i,i) \notin \beta$, and two bags Q_i and Q_j are complete in $B^{(\alpha,\beta)}$ if $(i,j) \in \beta$ and are anticomplete if $(i,j) \notin \beta$.

We also define a *star-graph* $S^{(\alpha,\beta)}$ with vertices $\{s_i : i \in \mathbb{N}\}$ and with $s_i s_j$ an edge whenever Q_i and Q_j are complete. This can be seen in $B^{(\alpha,\beta)}$ as an induced subgraph where we choose any single vertex from each Q_i.

Definition 5.1. The graph $H^{(\alpha,\beta)} = (V_P, E_P \triangle E_B)$ is called an *infinite path-clique graph* (i.e. the graph whose edges are the symmetric difference between those of the class-path P and the biclique graph $B^{(\alpha,\beta)}$). The corresponding *path-clique graph class* $\mathcal{H}^{(\alpha,\beta)}$ is the finite induced subgraphs of $H^{(\alpha,\beta)}$.

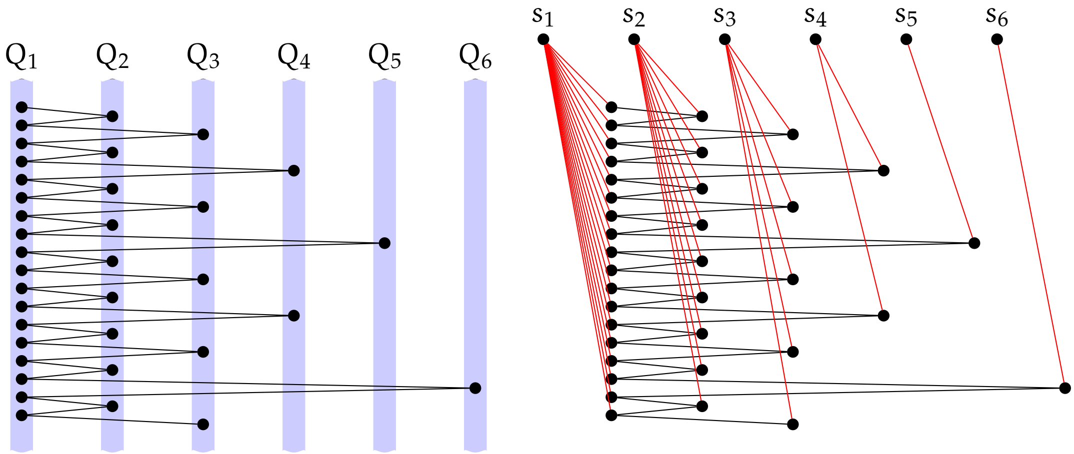


Figure 5.1: Comparing the path-clique graph $H^{(\alpha,\beta)}$ 'power graphs' (left) and the path-star graph R^{α} (right) where $\alpha = \kappa(2)$ and $\beta = \{(i,i) : i \in \mathbb{N}\}$ (path edges in black, clique edges in blue shaded areas, star-edges in red).

As an example, if $\alpha = \kappa(2)$ (defined in Section 4.2.3) and $\beta = \{(i,i) : i \in \mathbb{N}\}$ then $\mathcal{H}^{(\alpha,\beta)}$ is a path-clique graph class called the *power graphs* (see [49]) and $H^{(\alpha,\beta)}$ is shown on the left in Figure 5.1.

If α and β are clear from the context then we shorten $B^{(\alpha,\beta)}$, $S^{(\alpha,\beta)}$, $H^{(\alpha,\beta)}$ and $\mathcal{H}^{(\alpha,\beta)}$ to B, S, H and $\mathcal{H}$ respectively.

Any graph $G \in \mathcal{H}^{(\alpha,\beta)}$ can be witnessed by an embedding $\phi(G)$ into the infinite graph $H^{(\alpha,\beta)}$. To simplify the presentation we will associate G with a particular embedding in $H^{(\alpha,\beta)}$ depending on the context. We will be especially interested in subgraphs of G induced by the intersection of G with intervals of class-path P: in particular, an induced subgraph of G on the vertices of $V[G] \cap I_{[j,j+k]}$ will be denoted $G_{[j,j+k]}$.

5.2.2 Sparsification of a path-clique to a path-star graph

The transformation of a dense path-clique graph into a graph containing a sparse t-sail can be achieved using vertex-minor 'sparsification', that is, using a sequence of local complementations and vertex removals (recall the definition of local complementation in Section 1.3.3).

Notice that given a clique $K_n = (V_K, E_K)$ with a vertex $v \in V_K$, the vertex-minor $K_n * v$ is a star S_{n-1} with star-vertex v. Similarly, given a biclique $K_{m,n}$ with independent vertex sets U_K and V_K, with $u \in U_K$ and $v \in V_K$, the vertex-minor $G * v * u * v - v$ is a star S_{m-1} with star vertex u and an independent set of $n-1$ vertices.

If $G = B[V_G]$ is a finite induced subgraph of B on vertex set $V_G \subseteq V_P$, and $\mathsf{support}(G) = \{i : V_G \cap Q_i \neq \emptyset\}$ we define $S_G = S[\{s_i : i \in \mathsf{support}(G)\}]$. We then choose a set of *star-vertices* $V_S \subseteq V_G$ with precisely one vertex from each bag Q_i for $i \in \mathsf{support}(G)$ (the selection of vertices for V_S is dealt with later in Lemma 5.8). We re-label the vertices of V_S to $\{\overline{s_i}\}$ where $\overline{s_i} \in V_S$ corresponds to s_i where $i \in \mathsf{support}(G)$. We transform G by applying a sequence of local complementations in a defined order to the vertices in V_S to derive a locally equivalent graph G'.

We define four states for each bag Q_i in G' where $i \in \mathsf{support}(G)$:

1. a *clique* if $G'[Q_i]$ is a clique,
2. a *star* if $N(v) = \{\overline{s_i}\}$ for any vertex $v \in Q_i \setminus \{\overline{s_i}\}$,
3. a *link* if $G'[Q_i]$ is an independent set and Q_i is complete with some other bag Q_j for $j \in \mathsf{support}(G)$,
4. a *sink* if $G'[Q_i]$ is an independent set and $Q_i \setminus \{\overline{s_i}\}$ is anticomplete with all other sets $Q_j \setminus \{\overline{s_j}\}$ for $j \in \mathsf{support}(G)$. We refer to a sink where every vertex is of degree k in G' as a k-*sink*.

Example 5.2. *Suppose G is a graph with four bags Q_1, Q_2, Q_3 and Q_4, with star-vertices $\overline{s_1}$, $\overline{s_2}$, $\overline{s_3}$ and $\overline{s_4}$ respectively, where each bag is a clique and Q_1 is complete with each of Q_1, Q_2 and Q_3 (so S_G is a star S_3). If $G' = G * \overline{s_2} * \overline{s_3} * \overline{s_4}$ then it can be seen that Q_2, Q_3 and Q_4 are stars and Q_1 is a 3-sink in G'.*

Lemma 5.3. *If $G = B[V_G]$ is a finite connected induced subgraph of B, and G' the graph formed by any sequence of local complementations applied to G in the way described, then for all $i \in \mathsf{support}(G)$ the bag Q_i must be in one of these four states in G'.*

Proof. Local complementation on the star-vertex in Q_i, that is $\{\overline{s_i}\}$, transforms Q_i from a clique to a star, a star to a clique or does not change the state of Q_i if it is a link or sink.

Local complementation on a star-vertex from another bag, say $\{\overline{s_j}\}$ from Q_j, either has no impact on the state of Q_i, or transforms it from a clique to a link or sink, or transforms it from a link or sink to a clique. □

We say the graph G' is *semi-sparsified* if for every $i \in \mathsf{support}(G)$ the bag Q_i is either a star or a k-sink, and a semi-sparsified G' is *sparsified* if it has no k-sinks for $k > 1$.

Lemma 5.4. *Any biclique graph B can be semi-sparsified.*

Proof. From Lemma 5.3 we only have four bag states to consider. We demonstrate that we can construct a process consisting of a sequence of local complementations to systematically transform all clique and link bags to stars or sinks which proves the lemma.

First, note that if Q_i is a star it will remain a star so long as there is no further local complementation on vertex $\overline{s_i}$, so we only allow local complementations on the star-vertices of non-star bags. Thus, the number of stars can only increase with each local complementation.

Secondly, if Q_i is a clique it can always be transformed to a star by local complementation on its own star-vertex $\overline{s_i}$.

Lastly, if Q_i is a link then there exists another bag Q_j that itself is either a clique or link, such that Q_i is complete with Q_j (Q_j cannot be a star or sink as this contradicts the definitions). Local complementation on $\overline{s_j}$ followed by local complementation on $\overline{s_i}$ transforms link Q_i into a star.

We can now devise a process consisting of a series of steps, where at each step we select a remaining clique or link and transform it to a star, increasing the number of stars by at least one. Thus, the process must terminate. □

Note that semi-sparsification is not a unique process. It is possible that some sequences of local complementation are 'better' than others i.e., they result in more stars and fewer sinks.

k-sinks for $k > 1$ represent a blockage to full sparsification, since they are typically not sparse objects, and it is not clear how these could be removed or converted to stars. We need to look at particular star-graph structures to make progress.

First we deal with the case where S contains a number of non-trivial components (a vertex or component in S is trivial if it is an unconnected vertex corresponding to an isolated independent set in B):

Lemma 5.5. *If G is an induced subgraph of B with star-graph S_G that has t non-trivial components then G has a vertex-minor G' that is* sparsified *with t disjoint stars and, possibly, some isolated vertices.*

Proof. As each component is non-trivial it must contain at least one bag that is a clique or a link. Hence, once a component is semi-sparsified using the process in Lemma 5.4 it must contain at least one bag that is a star.

Other than this one star bag, delete all other star-vertices in the component and all that remains is one star and, possibly, some isolated vertices. Note that any sinks remaining are 1-sinks and therefore represent leaves of the star. Repeat for each of the t components. □

We now consider a star-graph that is a path:

Lemma 5.6. *If G is an induced subgraph of B with star-graph S_G that is a path with i edges then G has a locally equivalent G' that is* sparsified *with $i+1$ stars or i stars and one sink.*

Proof. Without loss of generality, let the vertices of S_G be $s_1, \dots s_{i+1}$ where $s_{k-1} \sim s_k \sim s_{k+1}$ for $2 \leqslant k \leqslant i$ with corresponding vertex set V_S of G. We apply local complementations to the star-vertices using the following iterative process to derive a locally equivalent graph G':

1. Set counter $k = 0$ and set $G' = G$;
2. Set $k = k + 1$;
3. If $k > i + 1$ then stop; otherwise:
   (a) If bag Q_k in G' is a clique, set $G' = G' * \overline{s_k}$;
   (b) Otherwise:
      i. Set $k = k + 1$;
      ii. If $k > i + 1$ then stop; otherwise:
         A. If bag Q_k in G' is a clique, set $G' = G' * \overline{s_k} * \overline{s_{k-1}}$;
         B. Otherwise, set $G' = G' * \overline{s_{k-1}} * \overline{s_k} * \overline{s_{k-1}}$;
4. Return to step 2;

This process must end when $k = i + 2$. The sparsification process turns every bag into a star. Once a star, a bag cannot revert back to a clique or link since we carry out no further local complementations on already sparsified bags. At the point there are only two non-sparsified bags left, if they are two cliques then the bag Q_{i+1} ends as a sink. Otherwise, the sparsification ends with no sinks. □

We can extend this process to trees:

Lemma 5.7. *If G is an induced subgraph of B and S_G is a tree with n vertices including q leaves then G has a locally equivalent G' that is semi-sparsified with at least $n - q + 1$ stars and at most $q - 1$ sinks, each of which is formed from a bag Q_i corresponding to a leaf of S_G.*

Proof. We first observe that as S_G is a tree it can be decomposed into $q-1$ disjoint paths P_k such that each P_k is an induced subgraph of S_G and $V[S_G] = \bigcup_k V[P_k]$. The first path is chosen as any induced path of S_G that has endvertices at two leaves. The second path is chosen that has one endvertex adjacent to a vertex of the first path and the other endvertex a leaf. The process continues with construction of further paths with one endvertex adjacent to one of the paths already constructed and the other endvertex a leaf. The process must finish when all q leaves of S_G are absorbed in one of the $q-1$ paths.

Let G_k be the subgraph of G induced by the vertices of bags Q_i corresponding to the path P_k in S_G. We now apply the process of Lemma 5.6 to each of the subgraphs G_k in turn. P_1 starts and ends at a leaf of S_G so the direction of sparsification does not matter. For subsequent paths P_k the sparsification starts at the endvertex of P_k that is internal to S_G and finishes at a leaf of S_G.

Notice that sparsification of G_1 has an impact on the neighbourhood of only the nearest bag of G_2 (corresponding to the internal endvertex of P_2). No loops have been created in P_2 so sparsification can proceed as in Lemma 5.6, and similarly for subsequent paths. □

As can be seen, even for paths and trees, full sparsification is not necessarily possible in all cases.

Sparsification so far has been described as to its effect on a finite graph $G = B[V_G]$ in B. However, we intend to apply sparsification to the corresponding path-clique graph $H[V_G]$ to transform it so that it contains a path-star graph (t-sail). We need to be careful that the star-vertices V_S used for local complementation are chosen so that we can establish the path components required for a t-sail. The following Lemma tells us that if we apply the same sparsification to $H[V_G]$ as was applied to sparsify $B[V_G]$ and exclude the star-vertices and the vertices either side of them in path P, this results in the class-path P remaining intact in the remaining graph, which can then be used for the path components required for a t-sail.

Lemma 5.8. *Suppose $G = B[V_G]$ is a finite induced subgraph of B with star-vertices V_S selected so that they form an independent set in P. Define $V_S^+ = \{p_i \in V_G : p_{i-1}, p_i,$ or $p_{i+1} \in V_S\}$ (i.e the star vertices and the vertices either side of them in path P).*

Let G be sparsified to G' by local complementation sequence L. Then

$$H * L[V_G \setminus V_S^+] = P[V_G \setminus V_S^+].$$

Proof. Suppose $u, v \in V \setminus V_S^+$ and $uv \in E_P$. If $uv \in E_B$ then from the symmetric difference requirement $uv \notin E[H]$. Sparsification complements the edge sets of cliques and complete sets in $B[V \setminus V_S^+]$ so $uv \in E[H * L]$. On the other hand, if $uv \notin E_B$ then from the symmetric difference requirement $uv \in E[H]$. Sparsification does not alter the independent sets and anticomplete sets in $B[V \setminus V_S^+]$ so $uv \in E[H * L]$. Thus, $uv \in E[H * L]$ if $uv \in E_P$.

A similar argument applies to show $uv \notin E[H * L]$ if $uv \notin E_P$. □

The graph $H * L[(V_G \setminus V_S^+) \cup V_S]$ is now an object approaching a t-sail since we have the necessary path components $P[V_G \setminus V_S^+]$ and star-vertices V_S. However, k-sinks for $k > 1$ are a problem as already mentioned and are obstructions to t-sail vertex-minors. Also, sparsification may generate unwanted edges between the star vertices.

To create a t-sail we need to show that we can choose star-vertices in a suitable way to create the necessary path components, and that we can avoid k-sinks for $k > 1$ and deal with edges between the star-vertices. These issues are addressed in the next section.

5.3 Path-clique graph classes with unbounded clique-width

By using sparsification we show that the class of power graphs is a member of a wider family of path-clique graph classes with unbounded clique-width.

To create sparsified t-sails we require an unbounded number of stars in G'. The usefulness of sparsification depends on the position of k-sinks for $k > 1$ as any pair of star-vertices that are connected to the vertices of a sink cannot both be used to form a t-sail. It would be nice to be able to choose the location of the sinks, but in practice this is hard (if not impossible), since the impact of local complementation on any type of irregular star graph itself becomes complicated. We consider two cases where we can avoid k-sinks:

Lemma 5.9. *Suppose $H^{(\alpha,\beta)}$ is an infinite path-clique graph, with auxiliary biclique graph B and star-graph S with at least $t \in \mathbb{N}$ non-trivial components. Then the path-clique class $\mathcal{H}^{(\alpha,\beta)}$ contains a graph containing a t-sail as a vertex-minor.*

Proof. From Lemma 5.5 we can sparsify B to give us a vertex-minor consisting of one star corresponding to each non-trivial component of S. Without loss of generality let these stars correspond to bags $Q_1, \ldots, Q_t$.

Identify disjoint intervals $I_1, \dots, I_t$ each of which includes a representative vertex from each bag Q_i for $1 \leqslant i \leqslant t$. This is always possible because every letter in α appears an unbounded number of times. Define a further disjoint interval I_{t+1} that contains at least one representative vertex from every bag Q_i that is represented in $\bigcup_{i=1}^{t} I_i$. Suppose there are $x \geqslant t$ such representatives. We label these representative vertices $V_S = \{\overline{s_i} \in Q_i : 1 \leqslant i \leqslant x\}$. The vertices of V_S are chosen such that they form an independent set in P, which is always possible since if two of these vertices are adjacent we just extend the interval until we find another vertex in the desired set. We use this set of vertices V_S as the star-vertices for the local complementation.

Set $G = B\left[\bigcup_{i=1}^{t+1} I_i\right]$ then applying Lemma 5.5 to each component there is a sequence of local complementations on the set V_S, say L, that sparsifies G to G′ leaving each bag $Q_1, \dots, Q_t$ as a disjoint star in G′.

Let $H = H^{(\alpha,\beta)}\left[\bigcup_{i=1}^{t+1} I_i\right]$, then $H' = (H * L)\left[\bigcup_{i=1}^{t} I_i \cup \{\overline{s_1}, \dots, \overline{s_t}\}\right]$ is a vertex-minor of a graph in $\mathcal{H}^{(\alpha,\beta)}$. From Lemma 5.8 the sparsification results in the class-path component $P[I_i]$ for $1 \leqslant i \leqslant t$ being an induced subgraph of H′. Furthermore, H′ contains t disjoint star-vertices $\overline{s_1}, \dots, \overline{s_t}$ whose leaves embed in each $P[I_i]$ for $1 \leqslant i \leqslant t$ and the other vertices in each path component can only be degree two as the other stars have been removed. This satisfies the requirement for H′ to contain a t-sail. □

To overcome the problem with k-sinks in trees, in the next lemma we use the idea of a *rooted leafless tree* which is an infinite tree with only one vertex of degree one (the root). The branches of the tree consequently each contain an unbounded number of vertices. We also require the classic Ramsey result, Theorem 2.10.

Definition 5.10. A *clique* t*-sail* is a t-sail with additional edges so that the t star-vertices form a clique and not an independent set.

Lemma 5.11. *Suppose $H^{(\alpha,\beta)}$ is an infinite path-clique graph, with auxiliary biclique graph B and star-graph S that contains a component that is a rooted leafless tree. Then for any $t \in \mathbb{N}$, the path-clique class $\mathcal{H}^{(\alpha,\beta)}$ contains a graph with a t-sail or a clique t-sail vertex-minor.*

Proof. Let $x = 2^{2t-3}$. Select x bags Q_i from the component of B corresponding to the rooted leafless tree in S. Without loss of generality let these be bags $Q_1, \dots, Q_x$.

As in the previous Lemma, identify disjoint intervals $I_1, \dots, I_x$ each of which includes a representative vertex from each bag Q_i for $1 \leqslant i \leqslant x$. Suppose there are $y \geqslant x$ bags Q_i represented in $\bigcup_{i=1}^{x} I_i$. Without loss of generality let these be bags $Q_1, \dots, Q_y$.

As before, define a further disjoint interval I_{x+1} that contains at least one representative vertex from every bag $Q_1, \ldots, Q_y$ which form an independent set. We label these representative vertices $V_S = \{\overline{s_i} \in Q_i : 1 \leqslant i \leqslant y\}$, with corresponding vertices $\{s_1, \ldots, s_y\}$ in S.

From the rooted leafless tree in S, for each $i = 1, \ldots, y$ find the next vertex that sits on the same branch as s_i away from the root that is not already in $\{s_1, \ldots, s_y\}$. This is always possible as the tree is leafless. Without loss of generality let this set be $\{s_{y+1}, \ldots, s_{2y}\}$. We extend interval I_{x+1} to include a representative from each bag $Q_{y+1}, \ldots, Q_{2y}$, which we call $V_Y = \{\overline{s_{y+1}}, \ldots \overline{s_{2y}}\}$.

Set $G = B\left[\bigcup_{i=1}^{x+1} I_i\right]$ then S_G is a tree with leaves in $\{s_{y+1}, \ldots s_{2y}\}$. So from Lemma 5.7 there is a sequence of local complementations on the set $V_S \cup V_Y$, say L, that sparsifies G to G' leaving each bag $Q_1, \ldots, Q_y$ as a disjoint star in G' (i.e. none of them are sinks).

Let $G_H = H^{(\alpha,\beta)}\left[\bigcup_{i=1}^{x+1} I_i\right]$, and $G'_H = (G_H * L)\left[\bigcup_{i=1}^{x} I_i \cup \{\overline{s_1}, \ldots, \overline{s_x}\}\right]$. From Lemma 5.8 the sparsification results in the class-path component $P[I_i]$ for $1 \leqslant i \leqslant x$ being an induced subgraph of G'_H. Furthermore, G'_H contains the x disjoint stars $s_1, \ldots, s_x$ whose leaves embed in each $P[I_i]$ for $1 \leqslant i \leqslant x$, together with some unspecified edges between the star-vertices. It follows that G'_H is an x-sail except that there are some unspecified edges between the star-vertices.

Using Ramsey's Theorem 2.10 we deduce that our x star-vertices must contain either an independent set of size t or a clique of size t. If an independent set, it follows that H contains a t-sail as a vertex-minor and if a clique then H contains a clique t-sail as a vertex-minor. □

Lemma 5.12. *A class of graphs containing a clique* t*-sail for all* $t \in \mathbb{N}$ *has unbounded clique-width.*

Proof. Let $\mathcal{C}$ be a class of graphs containing a clique t-sail T_t for all $t \in \mathbb{N}$.

Let T_t^+ be the graph T_t with an additional vertex v with neighbourhood all vertices in the clique of T_t, and $\mathcal{C}^+$ a graph class containing T_t^+ for all $t \in \mathbb{N}$. Define the graph $T_t^- = T_t^+ * v - v$, and $\mathcal{C}^-$ a graph class containing T_t^- for all $t \in \mathbb{N}$. Notice that the graph T_t^- is a t-sail.

A t-sail has tree-width at least $t-1$ so $\mathcal{C}^-$ has unbounded tree-width and, as the graphs are sparse, from Theorem 1.16 and Lemma 1.10, unbounded rank-width and clique-width. From Lemma 1.11 $rw(T_t^+) \geqslant rw(T_t^-)$ it follows that $\mathcal{C}^+$ has unbounded rank-width and clique-width. From Lemma 1.5 $cw(T_t) \geqslant \frac{1}{2}\, cw(T_t^+)$ and it follows that $\mathcal{C}$ has unbounded clique-width. □

Theorem 5.13. *A path-clique hereditary class of graphs $\mathcal{H}^{(\alpha,\beta)}$, where the auxiliary star-graph S contains an unbounded number of non-trivial components or a component that is a rooted leafless tree, has unbounded clique-width.*

Proof. If S is a forest with an unbounded number of non-trivial tree components or a forest containing a rooted leafless tree then from Lemmas 5.9 and 5.11 for any $t \in \mathbb{N}$, the path-clique class $\mathcal{H}^{(\alpha,\beta)}$ contains a graph containing a t-sail or a clique t-sail as a vertex-minor.

From Lemma 5.12 the vertex-minor closure of $\mathcal{H}^{(\alpha,\beta)}$ has unbounded clique-width and therefore from Lemma 1.11 $\mathcal{H}^{(\alpha,\beta)}$ itself has unbounded rank-width/clique-width. □

Beyond the cases where the auxiliary star-graph contains an unbounded number of non-trivial components or a component that is a rooted leafless tree we conjecture:

Conjecture 5.14. *A path-clique hereditary class of graphs $\mathcal{H}^{(\alpha,\beta)}$ has unbounded clique-width if and only if its auxiliary biclique graph contains an unbounded number of disjoint cliques or bicliques.*

We denote Γ as the set of (α, β)-pairs for which path-clique hereditary class of graphs $\mathcal{H}^{(\alpha,\beta)}$ has unbounded clique-width.

5.4 Minimal path-clique graph classes

We show that the class of power graphs is a member of a wider family of path-clique graph classes that are minimal of unbounded clique-width.

To show that for some $(\alpha, \beta) \in \Gamma$ the path-clique graph class $\mathcal{H}^{(\alpha,\beta)}$ is a minimal class of unbounded clique-width we must show that any proper hereditary subclass $\mathcal{C}$ has bounded clique-width. If $\mathcal{C}$ is a hereditary graph class such that $\mathcal{C} \subsetneq \mathcal{H}^{(\alpha,\beta)}$ then there must exist a non-trivial finite forbidden graph F that is in $\mathcal{H}^{(\alpha,\beta)}$ but not in $\mathcal{C}$. In turn, this graph F must be an induced subgraph of $H^{(\alpha,\beta)}_{[j,k]}$ for some interval $[j, k]$, and thus $\mathcal{C} \subseteq \text{Free}(H^{(\alpha,\beta)}_{[j,k]})$.

We know that for a minimal class, α must be recurrent, because if it contains a factor $\alpha_{[j,k]}$ that either does not repeat, or repeats only a finite number of times, then $\mathcal{H}^{(\alpha,\beta)}$ cannot be minimal, as forbidding the induced subgraph $H^{(\alpha,\beta)}_{[j,k]}$) would leave a proper subclass that still has unbounded clique-width. Therefore, we only need consider recurrent α.

5.4.1 Words with almost fixed support

In this section we focus on words α that are almost periodic with *almost fixed support*.

The *support* of a word α is the set of letters $\mathcal{A}^\alpha$ that appear in it. The word α is said to have almost fixed support if for any $n \in \mathbb{N}$ there exists a fixed subset $\mathcal{A}_f^\alpha(n) \subseteq \mathcal{A}^\alpha$ of size $\jmath(n)$ such that every factor of α of length n contains at most one letter not in $\mathcal{A}_f^\alpha(n)$. Clearly $\mathcal{A}_f^\alpha(n) \subseteq \mathcal{A}_f^\alpha(n+1)$ for all $n \in \mathbb{N}$.

To prove such words exist we give the following example:

Lemma 5.15. *Nested words with one letter branches have almost fixed support.*

Proof. Let $\alpha^{\mathcal{S}_1}$ be a nested word over alphabet $\mathcal{S}_1$ with one letter branches and base $\mathcal{B}_1$. Define subwords $\alpha^{\mathcal{S}_i}$ for $i \geqslant 2$ by $\mathcal{S}_i = \mathcal{S}_{i-1} \setminus \mathcal{B}_{i-1}$ with base $\mathcal{B}_i$. Let $\mathcal{A}_f^\alpha(n) = \cup_{i=1}^n \mathcal{B}_i$.

Given the definition of nested, two letters x and y not in $\mathcal{A}_f^\alpha(n)$ must be separated by at least one letter from each base $\mathcal{B}_i$ for $1 \leqslant i \leqslant n$. In other words, a factor of $\alpha^{\mathcal{S}_1}$ of length n contains at most one letter not in $\mathcal{A}_f^\alpha(n)$. Therefore $\alpha^{\mathcal{S}_1}$ has almost fixed support. □

In Section 4.2.3 there is a set of nested words described as 'power nested words', including q-ary and Fibonacci representation words, that have one letter branches and are almost periodic. Thus the power nested words are examples of words in this category.

5.4.2 When $\mathcal{H}^{(\alpha,\beta)}$ is a minimal class of unbounded clique-width

We first demonstrate a bound on the linear clique-width (and therefore clique-width) of a graph G in $\mathcal{H}^{(\alpha,\beta)}$ where G only has vertices from a fixed number of bags Q_i.

Lemma 5.16. *If $\mathcal{A}_n$ is a set of n letters and $G = (V, E)$ is a graph in $\mathcal{H}^{(\alpha,\beta)}$ such that $V \subset \{p_i : \alpha_i \in \mathcal{A}_n\}$ then* $lcw(G) \leqslant n + 2$.

Proof. We construct G along class-path P. Let the highest path index for a vertex in G be t. We create $n + 2$ labels, one label c for the current vertex, one l for the previous vertex in path P and n set labels $q_1, \ldots, q_n$ corresponding to the n different letters in $\mathcal{A}_n$, which without loss of generality we can write as $1, \ldots, n$.

We construct the graph G by applying the following iterative process (to a 'work-in-progress' graph G_0) that defines our linear clique-width expression τ:

1. Set G_0 as the null graph and set counter $i = 0$;

2. Set counter $i = i + 1$;

3. If $p_i \in V$ (vertex creation and path edges)

   (a) Construct 'current' vertex p_i and give it label c (operation $c(p_i)$);

   (b) Take the disjoint union of G_0 with p_i (operation $G_0 \oplus p_i$);

   (c) If $(\alpha_{i-1}, \alpha_i) \notin \beta$, construct an edge from vertex labelled c to vertex labelled l (operation $\eta_{c,l}$);

   (d) Set counter $j = 0$ (clique and biclique edges);

      i. Set counter $j = j + 1$;

      ii. If $(\alpha_i, j) \in \beta$ construct edges from vertices labelled c to those labelled q_j (operation η_{c,q_j});

      iii. If $j < n$ return to step 3a.

4. Relabel vertex l to q_j where $\alpha_{i-1} = j$ and c to l (operations $\rho_{l \to q_j}, \rho_{c \to l}$);

5. If $i = t$ stop, otherwise return to step 2.

This process creates a linear-clique width expression τ that creates G_0 isomorphic to our given graph G and can be represented as a caterpillar tree, tree(τ), whose leaves correspond to the operations of vertex creation (step 4), using at most $n + 2$ labels. □

We now extend this idea to a bound on the linear clique-width (and therefore clique-width) of a path-clique graph where α has almost fixed support.

Lemma 5.17. *If* $G = (V, E)$ *is a graph in* $\mathcal{B}$ *with corresponding star-graph* S_G *then*

$$lcw(S_G) \leqslant lcw(G) \leqslant lcw(S_G) + 2.$$

Proof. $lcw(S_G) \leqslant lcw(G)$ since S_G is an induced subgraph of G.

Suppose τ_{S_G} is a linear clique-width expression for S_G requiring k labels, with caterpillar tree(τ_{S_G}). We can extend tree(τ_{S_G}) to a caterpillar tree(τ_G) with the use of only two additional labels. Each time there is a leaf in tree(τ_{S_G}) representing the creation of a vertex s_i we replace this in tree(τ_G) with a sub-caterpillar creating every vertex in bag Q_i. This only requires at most two additional labels since Q_i is only a clique or independent set. All vertices in Q_i have the same adjacencies outside Q_i so are all given the same label as s_i, leaving the two extra labels available for the next set. Constructed in this way τ_G is a linear clique-width expression for G requiring only $k + 2$ labels, so $lcw(G) \leqslant lcw(S_G) + 2$. □

Hence, the biclique graph B has bounded linear clique-width if and only if the corresponding star-graph S has bounded linear clique-width.

Lemma 5.18. *Let (α, β) be a pair such that*

(i) α has almost fixed support, so that every factor of α of length n has at most one letter outside a fixed alphabet of size $\jmath(n)$, and

(ii) the auxiliary star-graph S has bounded linear clique-width k.

Then for any graph $G \in H^{(\alpha,\beta)}$ such that the length of the longest interval in G is n, the linear clique-width of G is at most $k2^{\jmath(n)} + \jmath(n) + 2$.

Proof. Let α have support $\mathcal{A}^\alpha$ and the support for a factor of α of length n be $\mathcal{A}^\alpha_f(n)$ of size $\jmath(n)$ plus one other letter. Without loss of generality let $\mathcal{A}^\alpha_f(n) = \{1, \ldots, \jmath(n)\}$. For a graph $G \in H^{(\alpha,\beta)}$ with longest interval n, let $\mathcal{A}_\nu(G) = \{i \in \mathcal{A}^\alpha : Q_i \cap V(G) \neq \emptyset, i \notin \mathcal{A}^\alpha_f(n)\}$.

Firstly, consider construction of the subgraph G_ν of G induced by those vertices $p_z \in V(G)$ where $\alpha_z = i$ and $i \in \mathcal{A}_\nu(G)$. These vertices are an independent set in the class-path P (since there can only be one of them in every interval of length n) so G_ν is an induced subgraph of B. As S_{G_ν} has bounded linear clique-width k, from Lemma 5.17 it is possible to construct a linear clique-width expression τ_ν for G_ν with at most $k+2$ labels.

Secondly, consider construction of a subgraph G_j of G induced by an interval I_j in $V(G)$. Each G_j has a support of at most $\jmath(n)+1$ letters so from Lemma 5.16 there exists a linear clique-width expression τ_j for G_j with at most $\jmath(n)+3$ labels.

Our construction of a linear clique-width expression for G combines the expression τ_ν for G_ν with the expressions τ_j for each G_j.

We use the following labels:

(i) A current vertex label c and a last vertex label l;

(ii) $\jmath(n)$ labels $q_1, \ldots, q_{\jmath(n)}$ for the fixed support labels for vertices corresponding to the letters in $\mathcal{A}^\alpha_f(n)$;

(iii) $k2^{\jmath(n)}$ labels (x, y) for vertices corresponding to the letters in $\mathcal{A}_\nu(G)$: letter x is one of up to k labels potentially required for the linear clique-width expression τ_ν and letter y is one of up to $2^{\jmath(n)}$ labels representing all the possible combinations of adjacencies to the set of vertices with the fixed support set of labels $q_1, \ldots, q_{\jmath(n)}$.

Caterpillar tree(τ_ν) represents the linear clique-width expression defining G_ν. Each leaf of the caterpillar tree represents the creation of a vertex $p_z \in V(G_\nu)$. There is at most one such vertex in each interval of $V(G)$. The caterpillar imposes a linear order on the construction of vertices in G_ν. We use this ordering to construct the intervals of $V(G)$.

First we construct the intervals that only contain letters from $\mathcal{A}_f^\alpha(n)$. From Lemma 5.16 we can do this using only letters c, l and $q_1, \dots, q_{\jmath(n)}$, leaving each vertex in set Q_i with label q_i.

We construct the rest of our clique-width expression using the order from caterpillar tree(τ_ν) but replacing the construction of vertex $p_z \in V(G_\nu)$ with the construction of interval $I_j \in V(G)$ being the interval containing vertex p_z. We have already established that such an interval can be constructed with at most $\jmath(n)+3$ labels. However, we must be able to connect the vertex p_z to subsequently constructed intervals. For this we need labels (x, y) [label set (iii)] representing all the possible combinations of adjacencies to vertices from bags $Q_1, \dots Q_{\jmath(n)}$.

Vertices are constructed with each new vertex labelled 'c', edges added to other previously constructed vertices, including the path edge to the vertex labelled l. The vertex labelled l is subsequently allocated another label to identify it for future purposes, as follows. The labels $q_1, \dots, q_{\jmath(n)}$ are used for vertices corresponding to the letters in $\mathcal{A}_f^\alpha(n)$ as in Lemma 5.16. Once allocated these labels never change. The current vertex labelled c is relabelled l.

For vertices corresponding to a letter in $\mathcal{A}_\nu(G)$ we use label set (iii) of form (x, y). Letter x is the first label allocated to the vertex of the k labels used in tree(τ_ν) and y is one of $2^{\jmath(n)}$ possible labels signifying the neighbourhood of this vertex to the 'fixed support' vertex sets corresponding to the $\jmath(n)$ letters in $\mathcal{A}_f^\alpha(n)$. The first label x may change again later in the construction, determined by the linear clique-width expression τ_ν so that a label (x_1, y) may change to, say, (x_2, y). However, the second letter in the pair, y, does not change since this determines the (unchanging) neighbourhood in the 'fixed support' vertex sets.

It can now be seen that this caterpiller and sub-caterpillar structure ensures that all edges between the vertex sets corresponding to the two letter sets $\mathcal{A}_f^\alpha(n)$ and $\mathcal{A}_\nu(G)$ can be constructed, and τ_G is a linear clique-width expression for G using at most $k2^{\jmath(n)} + \jmath(n) + 2$ labels. □

We define a subset $\Gamma_{min} \subset \Gamma$ such that if $(\alpha, \beta) \in \Gamma_{min}$ the hereditary path-clique graph class $\mathcal{H}^{(\alpha,\beta)}$ is minimal of unbounded clique-width.

Theorem 5.19. *If* (α, β) *is a pair such that* *(i)* $(\alpha, \beta) \in \Gamma$,

(ii) α *is almost periodic and has almost fixed support, and*

(iii) *the auxiliary star-graph* $S^{(\alpha,\beta)}$ *has linear clique-width bounded by* k,

then $(\alpha, \beta) \in \Gamma_{min}$.

Proof. Let the word α have support $\mathcal{A}^\alpha$ and $\alpha_{[x,y]}$ be a factor of α for some $x < y \in \mathbb{N}$. As α is almost periodic there exists a constant $\ell(\alpha_{[x,y]})$ such that every factor of α of length at least $\ell(\alpha_{[x,y]})$ contains $\alpha_{[x,y]}$ as a factor. Furthermore, as α has almost fixed support, for any $n \in \mathbb{N}$ there exists a fixed subset $\mathcal{A}^\alpha_f(n) \subseteq \mathcal{A}^\alpha$ of size $\jmath(n)$ such that every factor of α of length n contains at most one letter not in $\mathcal{A}^\alpha_f(n)$.

Let $\mathcal{C}$ be a proper hereditary subclass of $\mathcal{H}^{(\alpha,\beta)}$ with forbidden graph $H^{(\alpha,\beta)}_{[x,y]}$. As α is almost periodic the subgraph of $H^{(\alpha,\beta)}$ induced by an interval of length at least $\ell(\alpha_{[x,y]})$ contains the forbidden graph. It follows that any graph $G \in \mathcal{C}$ cannot contain an interval of vertices longer than $\ell(\alpha_{[x,y]})$.

Lemma 5.18 now tells us that

$$lcw(G) \leqslant k2^{\jmath(\ell(\alpha_{[x,y]}))} + \jmath(\ell(\alpha_{[x,y]})) + 2,$$

and as k and $\jmath(\ell(\alpha_{[x,y]}))$ are fixed, $\mathcal{C}$ has both bounded linear clique-width and bounded clique-width. □

5.5 Uncountably many minimal classes

As for the first framework in Section 3.3, we show that there is an uncountably infinite number of path-clique minimal classes.

Lemma 5.20. *Let* γ *and* δ *be two distinct infinite binary words starting with* 1 *and each containing an infinite number of* 1*s. We denote* $\beta(\gamma) = \{(i,i) : \gamma_i = 1\}$ *and* $\beta(\delta) = \{(i,i) : \delta_i = 1\}$. *Then* $\mathcal{H}^{(\kappa(2),\beta(\gamma))}$ *and* $\mathcal{H}^{(\kappa(2),\beta(\delta))}$ *are distinct minimal classes.*

Proof. Let $n \geqslant 2$ be the first position in each word such that $\gamma_n \neq \delta_n$ (so $\gamma_i = \delta_i$ for $i < n$). Without loss of generality let $\gamma_n = 1$ and $\delta_n = 0$. We will show that the graph $G = H^{(\kappa(2),\beta(\gamma))}[2^{n-1}, 3 \times 2^{n-1}]$ cannot be embedded in $H^{(\kappa(2),\beta(\delta))}$ and hence the two classes are distinct.

Firstly, in $H^{(\kappa(2),\beta(\gamma))}$, alternate (odd) vertices of the path are in bag Q_1 which is the largest clique. The two path endvertices have only one edge to a vertex in Q_1, are the only two vertices in bag Q_n and are adjacent since $\gamma_n = 1$.

In $H^{(\kappa(2),\beta(\delta))}$ the largest clique in G can only be embedded in Q_1 since it contains alternate vertices of the path. This determines the endvertices. In an interval of length 2^n each bag Q_i for $i < n$ has more than two representatives, and for $i > n$ has at most one representative. Thus the two path endvertices can only be in Q_n. But $\delta_n = 0$ so the path endvertices are not adjacent - a contradiction. Thus G cannot embed in $H^{(\kappa(2),\beta(\delta))}$ and the two classes are distinct. □

Theorem 5.21. *There exists an uncountably infinite number of path-clique minimal hereditary classes of graphs of unbounded clique-width.*

Proof. There are an uncountable number of binary words with an infinite number of 1s starting with a 1. Each such word γ corresponds to a path-clique hereditary class $H^{(\kappa(2),\beta(\gamma))}$ which has auxiliary star-graph S containing an unbounded number of non-trivial components so by Theorem 5.13, $(\kappa(2), \beta(\gamma)) \in \Gamma$.

The word $\kappa(2)$ is almost periodic and from Lemma 5.15 has almost fixed support. Furthermore, S has linear clique-width of two, so by Theorem 5.19, $(\kappa(2), \beta(\gamma)) \in \Gamma_{min}$.

By Lemma 5.20, for any two such distinct words γ, the corresponding path-clique graph classes $\mathcal{H}^{(\kappa(2),\beta(\gamma))}$ are pairwise distinct, and hence we have an uncountably infinite number of path-clique minimal classes. □

5.5.1 Further thoughts

Sparsification provides a useful technique for proving certain dense hereditary classes have unbounded clique-width. Beyond the cases where the auxiliary star-graph is a forest with an unbounded number of non-trivial tree components or a forest containing a rooted leafless tree is more complex and not attempted here. The barriers to going further include the problem of sparsifying when the star-graph contains loops.

Theorem 5.19 gives us a new family of minimal hereditary graph classes of unbounded clique-width that are not covered by the grid framework in Chapters 2 and 3.

We now have two frameworks for minimal hereditary graph classes of unbounded clique-width, the framework in Chapters 2 and 3 based on a triple $\delta = (\alpha, \beta, \gamma)$ and the framework in this chapter based on a pair (α, β). It can be observed that there are strong similarities between the β, γ pair in the first framework and the β pairs in the second, given that they take a set of vertex sets or 'bags' and turn the bags into either cliques or independent sets, and create complete bipartite links between the bags. On the other hand, the α word in the first framework defines edge sets (including half-graphs) between consecutive columns in a grid but in the second α defines a single path determining which bag each vertex on the path goes in. Whether these two frameworks can be unified remains to be established.

There is also the interesting question: How does sparsification (by local complementation) turn a minimal dense hereditary graph class into a non-minimal (but nevertheless of unbounded clique-width) sparse class? An example of this is the minimal path-clique power graphs $\mathcal{H}^{(\kappa(2),\beta)}$ (where $\beta = \{(i,i) : i \in \mathbb{N}\}$) that sparsify into the path-star class $\mathcal{R}^{\kappa(2)}$ that does not contain a minimal subclass. Somehow sparsification can create an infinite antichain of unbounded clique-width.

Chapter 6

Antichains of unbounded clique-width

6.1 Introduction

We identify a further object that is an obstruction to bounded clique-width: for arbitrarily large t, a t-*clipper*.

We consider the class of *path-half* graphs which consists of the induced subgraphs of a half graph, where one part of the bipartition is replaced by a linear forest. This class was introduced by Korpelainen [42] (which he referred to as permutation-partition graphs) and has unbounded clique-width. We show that graphs in this class with large clique-width contain a large t-clipper as an induced subgraph, but that the class does not contain a minimal class of unbounded clique-width.

We give an example of a t-clipper antichain of unbounded clique-width and suggest other similar objects that will also be t-obstructions to bounded clique-width: a *clique* t-*clipper*, a *two-sided* t-*clipper* and *clique two-sided* t-*clipper*.

6.2 Path-half graphs

We start with the wider category of *path-bipartition* graphs before we define the specific class of *path-half graphs*:

Definition 6.1. A graph (V, E) is a *path-bipartition* graph if its vertices V can be partitioned into two sets V_P (path-vertices) and V_S (star-vertices) and its edges E can be partitioned into two sets, E_P (path-edges) and E_S (star-edges), so that

1. (V_P, E_P) is a forest of paths, and

2. (V, E_S) is a bipartition with independent sets, V_P and V_S.

Note that path-star graph classes (Definition 1.29) are a subset of path-bipartition graph classes.

Let $P_j = (V_j, E_j)$ be a path where $V_j = \{v_{j,i} : 1 \leqslant i \leqslant n\}$ and $E_j = \{v_{j,i}v_{j,i+1} : 1 \leqslant i \leqslant n-1\}$ for some $n \in \mathbb{N}$. Let $X_{j,t} = \{x_{j,1}, x_{j,2}, \dots, x_{j,t}\}$ be a set of t positive integers where $x_{j,1} < x_{j,2} < \cdots < x_{j,t}$ for each $1 \leqslant j \leqslant t$.

Next we define our new t-basic obstruction.

Definition 6.2. A path-bipartition graph $(V_P \cup V_S, E_P \cup E_S)$ is a t-*clipper* if the set V_S consists of t vertices, $s_1, \dots, s_t$, and there exist t sets $X_{j,t}$ as defined above, such that

1. The graph (V_P, E_P) comprises t path components $P_1, \dots, P_t$, and

2. $E_S = \{s_i v_{j,k} : 1 \leqslant i, j \leqslant t \text{ and } k \leqslant x_{j,i}\}$.

We can construct hereditary subclasses of path-bipartition graphs if we only allow certain types of edge set E_S between the vertex sets V_S and V_P. Korpelainen considered one such subclass, where the edges E_S form a half graph between the vertices of V_S and the vertices of each separate path component in V_P. He adopted the name *permutation-partition graphs* but we will use the name *path-half graphs* for consistency with other class names used here (see Figure 6.1) :

Definition 6.3. The hereditary class $\mathcal{H}$ of *path-half graphs* is the set of finite induced subgraphs of the infinite path-bipartition graph H whose vertices V can be partitioned into two sets V_P and V_S and whose edges E can be partitioned into two sets, E_P and E_S, where:

1. $V_P = \{v_{j,i} : j, i \in \mathbb{N}\}$,

2. $V_S = \{s_k : k \in \mathbb{N}\}$,

3. $E_P = \{v_{j,i}v_{j,i+1} : i, j \in \mathbb{N}\}$, and

4. $E_S = \{s_k v_{j,i} : 1 \leqslant i \leqslant k \in \mathbb{N}, j \in \mathbb{N}\}$.

We will call the vertices $v_{j,i}$ *path-vertices* and vertices s_k *star-vertices* where vertex $v_{j,i}$ sits in path (row) j and column i of H.

Observe that $\mathcal{H}$ contains a t-clipper for all $t \in \mathbb{N}$.

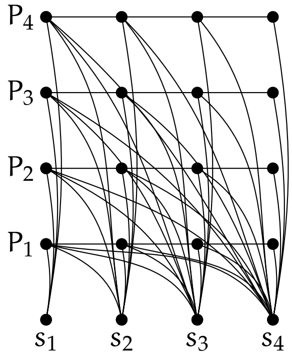


Figure 6.1: A 4-clipper and path-half graph

6.2.1 The graph class $\mathcal{H}$ has unbounded clique-width

In [42] $\mathcal{H}$ is shown to have unbounded clique-width using a square grid method. However, we want to focus on the slightly wider concept of a t-clipper, so we use a different method and provide a better bound.

Lemma 6.4. *The clique-width of a* t*-clipper is at least* $\lfloor \frac{t}{3} \rfloor$.

Proof. Let $G = (V, E)$ be a t-clipper.

As a reminder, any clique-width expression τ defining G can be represented as a rooted binary tree, tree(τ), whose leaves correspond to the operations of vertex creation, the internal nodes correspond to the $\oplus$-operation, and the root is associated with G. The operations η and ρ are assigned in the appropriate sequence along the respective edges of tree(τ). The tree is binary since each $\oplus$-operation brings together at most two previously constructed graphs. Also, it can be observed that an $\oplus$-vertex represents a subgraph of G on a vertex set $U \subset V$ but not usually an induced subgraph since there may still be edges to be created by η operations.

Let $\oplus_a$ be the lowest node in tree(τ) such that the constructed vertex-set, say U_a, contains either $\frac{2t}{3}$ star-vertices or $\frac{2t}{3}$ full paths. Suppose U_a contains $\frac{2t}{3}$ star-vertices. Letting the two children of $\oplus_a$ in tree(τ) be $\oplus_b$ and $\oplus_c$, then by assumption both U_b and U_c contain at most $\frac{2t}{3}$ full paths, but by the pigeonhole principle one of them must contain $\frac{t}{3}$ star-vertices. On the other hand, suppose U_a contains $\frac{2t}{3}$ full paths. Then by assumption both U_b and U_c contain at most $\frac{2t}{3}$ star-vertices, but by the pigeonhole principle one of them must contain $\frac{t}{3}$ full paths.

Hence, there must always be a node $\oplus_U$ such that either

1. Case 1: U contains at least $\frac{t}{3}$ of the star-vertices but at most $\frac{2t}{3}$ full paths, or

2. Case 2: U contains at most $\frac{2t}{3}$ star-vertices but at least $\frac{t}{3}$ full paths.

Let us colour vertices in U red and vertices in $V \setminus U$ white.

In Case 1, at node $\oplus_U$, we have at least $\frac{t}{3}$ paths not all red and at least $\frac{t}{3}$ red star-vertices, that we will denote as vertex set W. If at least one path is all white then the red vertices in W form a distinguished vertex set (defined in Section 2.3.1) so must have different labels. If no paths are all white then in each 'not all red' path there must be at least one red vertex adjacent to a white vertex. Let this set of red vertices be W'. Then W' is a set of at least $\frac{t}{3}$ vertices that are a distinguished vertex set and require different labels. Either way, with W or W', we need at least $\frac{t}{3}$ labels.

In Case 2, at node $\oplus_U$, we have at least $\frac{t}{3}$ white star-vertices, we denote set W, and at least $\frac{t}{3}$ all red paths. Taking one of the red paths, say P_j, let X be the vertices from the set $v_{j,1}, \dots, v_{j,t}$, as defined above, corresponding to the stars in W. Then the red vertices in X form a distinguished vertex set relative to white set W and we require at least $\frac{t}{3}$ labels.

Hence, any clique-width expression for G requires at least $\lfloor \frac{t}{3} \rfloor$ labels. □

As $\mathcal{H}$ contains a t-clipper for all $t \in \mathbb{N}$ it has unbounded clique-width.

6.2.2 Graphs in $\mathcal{H}$ with large clique-width contain a large t-clipper

Let $G = (V, E)$ be a finite graph in $\mathcal{H}$, and let vertices V be partitioned into sets V_P (path-vertices) and V_S (star-vertices) and edges E be partitioned into sets E_P and E_S, as previously described. Fix some embedding of G into the infinite graph H, and let $Q_j = \{i : v_{j,i} \in V_P\}$ (that is, the index of every vertex on path P_j for the given embedding of G) for all $j \in \mathbb{N}$ and $X = \{i : s_i \in V_S\}$ (that is, the index of every star-vertex for the given embedding of G).

Let $x_1 < \cdots < x_t$ be a set of t positive integers. We say the path P_j is $[x_1, x_t]$-*complete in* G if $Q_j \supset [x_1, x_{t-1} + 1]$ (that is, Q_j contains every integer between, and including, x_1 and $x_{t-1} + 1$, and therefore G contains an unbroken segment of path P_j between vertex v_{j,x_1} and vertex $v_{j,x_{t-1}+1}$). It can be observed that if $X \supset \{x_1, \dots, x_t\}$ and there are t or more paths P_j that are $[x_1, x_t]$-complete then G contains a t-clipper. We will use this fact to show that a subclass of $\mathcal{H}$ has unbounded clique-width if and only if it contains t-clippers for all t.

Theorem 6.5. *For every* $t \geqslant 1$, *every graph in* $\mathcal{H}$ *of (linear) clique-width greater than* $2t + 1$ *contains a* t*-clipper as an induced subgraph.*

Proof. For a contradiction, suppose there is a graph $G = (V, E) \in \mathcal{H}$ where $lcw(G) > 2t + 1$ that does not contain a t-clipper. We will show that we can construct a linear clique-width expression for G requiring at most $2t+1$ labels, thus showing that $lcw\ G \leqslant 2t + 1$, a contradiction.

Let the star-vertices s_i in G only have indices in $X = \{x_1, x_2, \ldots, x_m\}$ where $1 \leqslant x_1 < \cdots < x_m$, for some $m \in \mathbb{N}$. We know that, as G does not contain a t-clipper as an induced subgraph, we can take any t consecutive letters in X, say, $x_k, x_{k+1}, \ldots, x_{k+t-1}$, then at most $t-1$ of the P_j paths are $[k, k+t-1]$-complete in G.

We will use the following $2t + 1$ labels:

(i) 2 labels, c and p, for current and previous vertex for constructing new vertices;

(ii) $t-1$ labels, $l_1, \ldots, l_{t-1}$, for path-vertices with the same star neighbourhoods;

(iii) $t-1$ labels, $f_1, \ldots, f_{t-1}$, for the last vertex in unbroken paths.

(iv) 1 label, d, for 'dead' (all incident edges constructed) star-vertices.

We construct the graph G by applying the following iterative process (to a 'work-in-progress' graph G_0) that defines a clique-width expression τ. In overview, we are constructing G along the path segments from left to right, in stages, starting with path segments that begin to the left of the first star. At most $t-1$ of the paths are $[x_1, x_t]$-complete (i.e. at most $t-1$ of our path segments can continue beyond the $(t-1)$-th star), since otherwise we would have a t-clipper. We then construct the first star with corresponding edges. We can now recycle the labels and repeat the process for path segments that begin between the first and second stars, and then repeat, whilst continuing to recycling the $2t + 1$ labels.

1. Set G_0 as the null graph, path counter $j = 0$, star counter $i = 0$ and column counter $k = 0$;
2. NEXT STAR: Set $i = i + 1$;
3. NEXT PATH: Choose the next path $j \in [1, n]$ in the following order: first, those that have a vertex in column $x_{i+t-3} + 1$ with an l label, then other paths in any order.
4. If $i > 1$ set $k = x_{i-1}$;
5. NEXT VERTEX: Set $k = k + 1$;
6. If $k > \max(Q_j)$ go to [NEXT PATH].

7. (a) If $v_{j,k} \in V_P$ but not yet constructed then create the vertex and give it label c.
   (b) Otherwise,
      i. If $i < m$ and $k \geqslant x_i$ go to [NEXT PATH],
      ii. Otherwise go to [NEXT VERTEX].

8. Create edge from vertex labelled c to vertex labelled p;

9. Relabel vertex labelled p to its appropriate star neighbourhood label : that is, if $k-1 \in [1, x_1]$ relabel to l_1; if $k-1 \in [x_{i+y-2}+1, x_{i+y-1}]$ relabel to l_y, for $y \in [1, t-1]$;

10. (a) If $k = x_{i+t-2}+1$ then relabel vertex with label c to one of the unused f labels and go to [NEXT PATH] (Note, at most $t-1$ paths are $[x_i, x_i+t-1]$-complete so there are always enough f labels.);
      i. If $v_{j,k+1} \in V_P$ then relabel vertex labelled c to label p;
      ii. Otherwise, relabel vertex labelled c to its appropriate star neighbourhood l label;
    (b) Go to [NEXT VERTEX]

11. Create star-vertex s_{x_i} and give it label c;

12. create edge from vertex labelled c to vertices with label l_1;

13. Relabel vertex with label c to label d.

14. Relabel vertices with label l_j to 1_{j-1} for all $1 < j \leqslant t-1$ (frees up label l_{t-1}).

15. If $i < m$ go to [NEXT STAR];

This process ends when G_0 is isomorphic to our given graph G with clique width expression τ that requires at most $2t+1$ labels. □

6.2.3 The graph class $\mathcal{H}$ does not contain a minimal class

In [42] it is stated that $\mathcal{H}$ does not contain a minimal class but without proof. Here we prove that $\mathcal{H}$ does not contain a minimal subclass.

For $t \geqslant 3$ we define a sequences of graphs, $\{B_t\}$ in $\mathcal{H}$, as follows (see Figure 6.2):

$$B_t = H[\{v_{1,i} : 1 \leqslant i \leqslant t+2\} \cup \{s_1, s_{t+1}\}].$$

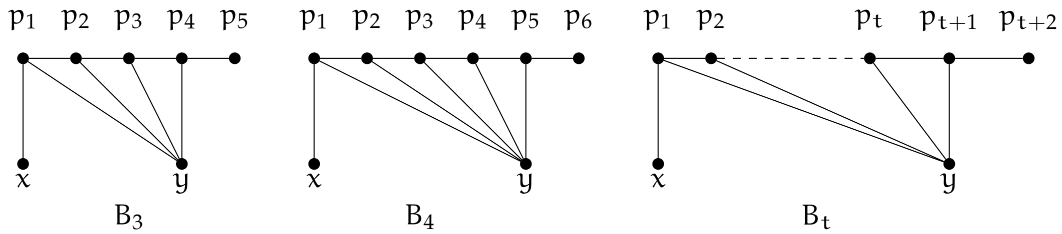
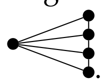


Figure 6.2: ’B’ antichain graphs B_3, B_4, B_t

Lemma 6.6. *The graph sequence* $B_3, B_4, \ldots$ *is an infinite antichain under the induced subgraph relationship.*

Proof. Suppose, for a contradiction, that B_i is an induced subgraph of B_j for $3 \leqslant i < j$ so there exists an embedding of B_i in B_j. The highest degree vertex (marked y in Figure 6.2) in B_i is the apex of a ‘fan’, drawn here: . The only possible embedding of y from B_i can be in the equivalent y in B_j. Vertices x and p_{i+2} in B_i are the only vertices not adjacent to y so must embed in vertices x and p_{j+2} in B_j (either way round). But the path xp_{i+2} in B_i is of length $i+2$ and the path xp_{j+2} in B_j of length $j+2$ and since $j > i$ the embedding is not possible. □

Lemma 6.7. *For integers* $t \geqslant 2$, $m > 3$ *and* $T \geqslant 2t$, *a* T*-clipper in* $\mathcal{H}$ *which does not contain any of* $B_1, \ldots, B_{m-1}$ *but does contain* B_m, *contains an induced subgraph isomorphic to a* t*-clipper which does not contain an induced subgraph* B_m.

Proof. Let G be a T-clipper in $\mathcal{H}$. Fix some embedding of G into H so that $G = (V_P \cup V_S, E_P \cup E_S)$ with the usual notation, and let $X = \{i : s_i \in V_S\}$ be the set of indices corresponding to the star-vertices V_S of G. Supposing the elements of X are listed in increasing order of size, let X' be the subset of X generated by taking the odd elements from this ordering with corresponding star-vertex set V'_S.

Let $G' = G[V_P \cup V'_S]$. We claim G' contains an induced subgraph isomorphic to a t-clipper which does not contain B_m. That G' contains an induced subgraph isomorphic to a t-clipper follows from the fact that it contains at least t of the star-vertices of G and all the path-vertices from the path components of G.

Suppose, for a contradiction, that G' contains an induced B_m.

Notice that for any graph in $\mathcal{H}$, any induced K_3 must have exactly two vertices in the same path of H and one star-vertex. Therefore, any embedding of B_m into G or G' must map the vertices of the path $p_1, \ldots, p_{m+1}$ from B_m into a single path of G or G' (see Figure 6.2). It follows that vertex y must map to a star-vertex and between x and

p_{m+2}, one must map to a star-vertex and one to a path-vertex for the same reasoning as in the previous lemma. Without loss of generality let x map to a star-vertex and p_{m+2} a path-vertex.

As B_m is the smallest such graph in G then x and y must be two consecutive star-vertices. But in G' we have removed alternate star-vertices so the path from p_1 to p_{m+1} in any possible embedding of B_m into G' must be of length at least $2m$, so B_m cannot embed in G', a contradiction. Hence G' does not contain an induced B_m. □

Theorem 6.8. *$\mathcal{H}$ does not contain a minimal class.*

Proof. If $\mathcal{D}$ is a minimal subclass of $\mathcal{H}$ then by Theorem 6.5, for every positive integer T, $\mathcal{D}$ contains a T-clipper.

Suppose the smallest 'B' graph in $\mathcal{D}$ is B_m $(m \geqslant 3)$. Then for $T \geqslant 2t$, from Lemma 6.7, $\mathcal{D}$ contains an induced subgraph isomorphic to a t-clipper which does not contain an induced subgraph B_m. Thus the subclass $\mathcal{D} \cap \mathrm{Free}(B_m)$ still contains a t-clipper for arbitrarily large t and has unbounded clique-width, which contradicts $\mathcal{D}$ being minimal. □

Let us define another sequence of graphs $\{A_t\}$ in $\mathcal{H}$ as follows:

$$A_t = H[\{v_{j,i} : 1 \leqslant i \leqslant (t-1)t+1\, , j \leqslant t, \} \cup \{s_k : k \in \{1, t+1, 2t+1, \ldots, (t-1)t+1\}\}]$$

.

It is easily observable that each A_t is a t-clipper and that $B_t \leqslant A_t$ for all $t \geqslant 3$. Equally, using the same arguments as used in the proof of Lemma 6.7, it is the case that $B_r \nleqslant A_s$, and therefore $A_r \nleqslant A_s$, for any $r < s$. Therefore, the sequence $A_3, A_4, \ldots$ is an antichain of unbounded clique-width (see Figure 6.3).

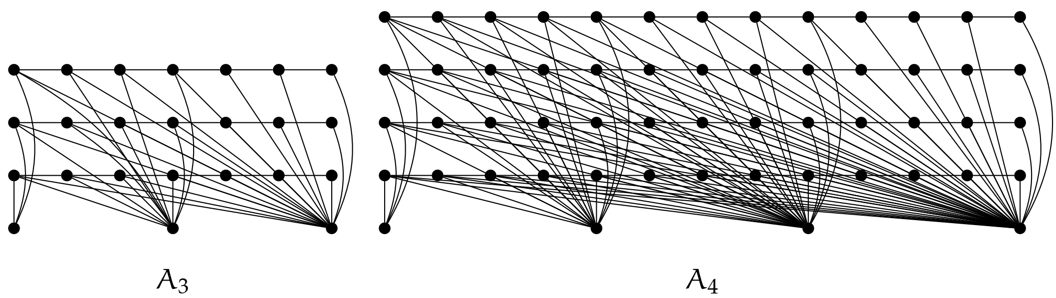


Figure 6.3: First two graphs A_3 and A_4 in an antichain of unbounded clique-width in $\mathcal{H}$

6.3 A family of antichains of unbounded clique-width

We have now seen the following examples of t-basic obstructions to bounded clique-width: for arbitrarily large t, a subdivision of a $t \times t$ wall, the line graph of a subdivision of a $t \times t$ wall, a t-sail and a t-clipper. Although we do not provide a formal proof, it is not too difficult to see that the methodology described here could be applied to similar obstructions - some of which we present in Figure 6.4: clique t-sails, clique t-clippers (a t-clipper in which the star-vertices are joined to form a clique), two-sided t-clippers (where the star-vertices are adjacent to all the path vertices except those with the same column index) and clique two-sided t-clippers. These all correspond to a family of antichains of unbounded clique-width.

These t-basic obstructions appear in pairs, where the stars are either an independent set or a clique. It would be a simple application of Ramsey Theory to show that, in fact, if there are any edges between the star-vertices, then given enough stars we could find an independent set or clique of arbitrarily large size. Similarly, it might be interesting to see whether, for any class consisting of stars and long paths, it is possible to create a characterization of unbounded clique-width using these t-basic obstructions.

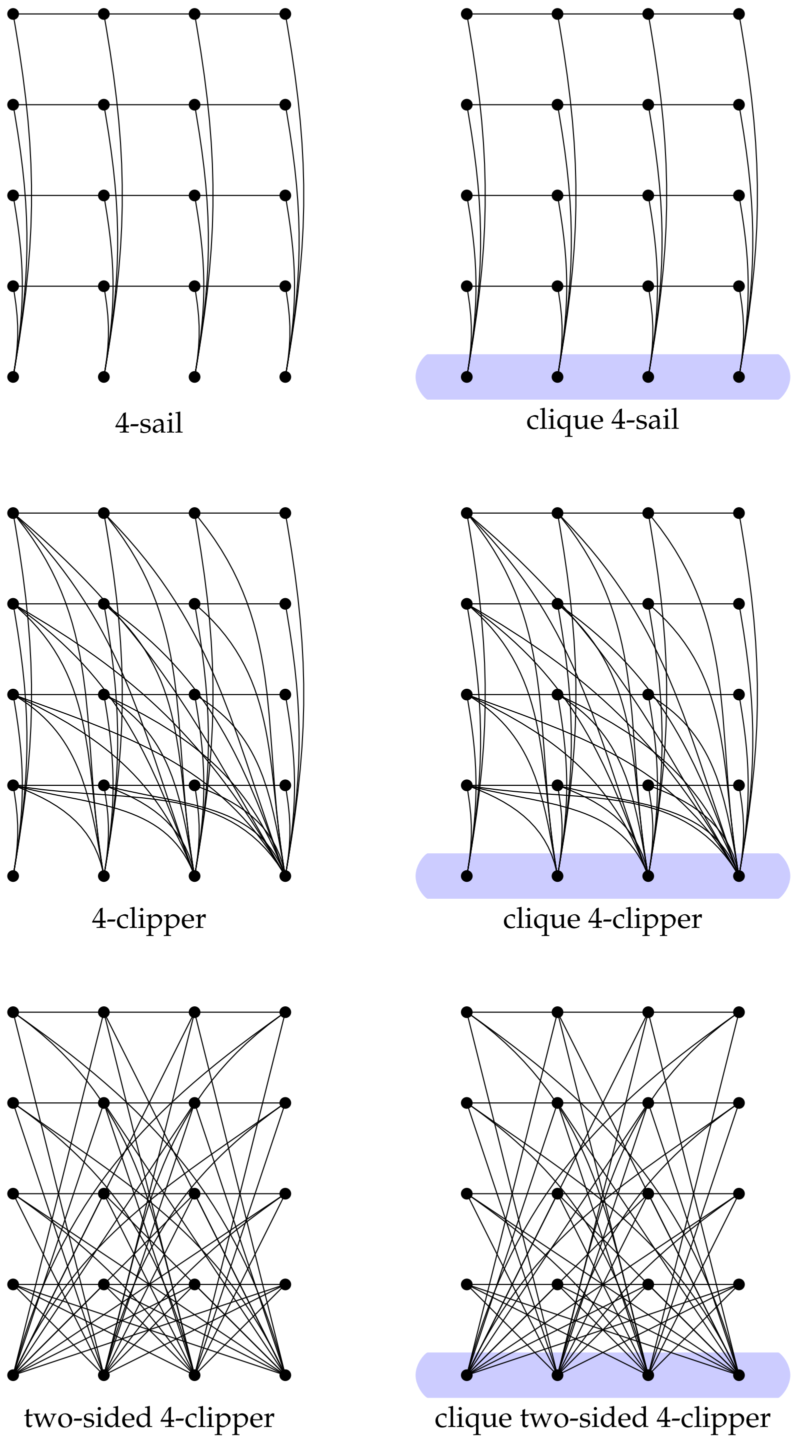


Figure 6.4: t-basic obstructions that block bounded clique-width (vertices in blue shaded area form a clique)

Chapter 7

Topics for further research

The ultimate goal is to characterize when a hereditary graph class has bounded clique-width, or in other words, characterize the obstructions to bounded clique-width. There are two types of such obstruction – minimal classes and antichains of unbounded clique-width.

This thesis provides a large collection of new minimal classes collected together in two frameworks together with some new t-basic obstructions to bounded clique-width that appear as antichains of unbounded clique-width. However, there still appears to be a long way to go before we could claim a full characterization. We suggest further areas for research in the following.

7.1 Dense classes

We now have two 'frameworks' for minimal classes, both of which contain uncountably many minimal classes.

The first framework described in Chapters 2 and 3 is largely complete, but the second framework in Chapter 5 still needs further development. We only assumed symmetric β edges between the vertex sets Q_i. If we allow non-symmetric β pairs (example, $(i,j) \in \beta$, $(j,i) \notin \beta$) this would introduce half-graph type structures into the framework. For example, if $(i,j) \in \beta$, $(j,i) \notin \beta$ then, rather than making bags Q_i and Q_j complete in the biclique graph, where $\alpha_m = i$ and $\alpha_n = j$ we could let $p_m \sim p_n$ if and only if $m < n$.

Conjecture 7.1. *Allowing non-symmetric β pairs into the framework will open up further examples of minimal hereditary graph classes of unbounded clique-width.*

Whether the two frameworks can be unified and whether there are other minimal classes not in either framework remain open questions.

As a consequence of the Grid Theorem for Vertex Minors [33] the circle graphs are the unique minimal vertex-minor-closed class of unbounded clique-width. Some of our minimal classes are subclasses of circle graphs (e.g. bipartite permutations) and some are not (e.g. power graphs). Exploring this dichotomy may offer one way of understanding the characteristics of minimal classes and the impact of local complementation.

Dawar and Sankaran [29] introduced the idea of using MSO transductions, showing that a class that 'interprets' arbitrarily large grids has unbounded clique-width. This technique may offer an alternative way to generate more minimal classes, and perhaps even show the way to prove that a complete characterization has been found.

The t-clipper and other t-basic obstructions revealed in Chapter 6 offer another direction to take research, identifying further such structures and using Ramsey theory to identify where such structures must exist.

7.2 Sparse classes

Although in Chapter 4 we added a new t-basic obstruction, a t-sail, the identification of a full list of such objects that are obstructions to bounded tree-width in sparse hereditary graph classes is still some way from being achieved. Our approach has been to consider graphs of bounded arboricity, in particular, those graphs of arboricity two constructed using forests of paths and stars. We believe this approach could be a fruitful way to identify further boundary objects i.e., by considering other objects of arboricity two such as 'star-star' classes and then progressing to objects of arboricity three etc.

In the sparse classes of unbounded tree-width studied here we have not identified any minimal classes of unbounded tree-width, which suggests Conjecture 1.31, that no sparse hereditary graph classes of unbounded tree-width contain a minimal class of unbounded tree-width. Proving this is one possible avenue for further investigation.

Using the recently discovered concepts of k-blocks [61] and strong k-blocks [3] may offer another way to explore sparse classes.

Bibliography

[1] P. Aboulker, I. Adler, E. J. Kim, N. L. D. Sintiari, and N. Trotignon. On the tree-width of even-hole-free graphs. *European J. Combin.*, 98:Paper No. 103394, 21, 2021.

[2] T. Abrishami, B. Alecu, M. Chudnovsky, S. Hajebi, and S. Spirkl. Induced subgraphs and tree decompositions VIII. Excluding a forest in (theta, prism)-free graphs, 2023. arXiv:2301.02138.

[3] T. Abrishami, B. Alecu, M. Chudnovsky, S. Hajebi, and S. Spirkl. Induced subgraphs and tree decompositions VII. Basic obstructions in H-free graphs. *J. Combin. Theory Ser. B*, 164:443–472, 2024.

[4] M. Albert, R. Brignall, N. Ruškuc, and V. Vatter. Rationality for subclasses of 321-avoiding permutations. *European Journal of Combinatorics*, 78:44–72, 2019.

[5] B. Alecu, M. Chudnovsky, S. Hajebi, and S. Spirkl. Induced subgraphs and tree decompositions XIII. Basic obstructions in $\mathcal{H}$-free graphs for finite $\mathcal{H}$, 2023. arXiv:2311.05066.

[6] B. Alecu, M. M. Kanté, V. V. Lozin, and V. Zamaraev. Between clique-width and linear clique-width of bipartite graphs. *Discrete Math.*, 343(8):111926, 14, 2020.

[7] V. E. Alekseev. On easy and hard hereditary classes of graphs with respect to the independent set problem. *Discrete Appl. Math.*, 132(1-3):17–26, 2003.

[8] A. Atminas, R. Brignall, V. V. Lozin, and J. Stacho. Minimal classes of graphs of unbounded clique-width defined by finitely many forbidden induced subgraphs. *Discrete Applied Mathematics*, 295:57–69, 2021.

[9] B. Bergougnoux, M. M. Kanté, and O.-J. Kwon. An optimal XP algorithm for Hamiltonian cycle on graphs of bounded clique-width. *Algorithmica*, 82(6):1654–1674, 2020.

[10] R. Boliac and V. V. Lozin. On the clique-width of graphs in hereditary classes. In *Algorithms and computation*, volume 2518 of *Lecture Notes in Comput. Sci.*, pages 44–54. Springer, Berlin, 2002.

[11] M. Bonamy, E. Bonnet, H. Déprés, L. Esperet, C. Geniet, C. Hilaire, S. Thomassé, and A. Wesolek. Sparse graphs with bounded induced cycle packing number have logarithmic treewidth. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3006–3028. SIAM, Philadelphia, PA, 2023.

[12] A. Brandstädt and V. V. Lozin. On the linear structure and clique-width of bipartite permutation graphs. *Ars Combin.*, 67:273–281, 2003.

[13] R. Brignall and D. Cocks. Uncountably many minimal hereditary classes of graphs of unbounded clique-width. *Electron. J. Combin.*, 29(1):Paper No. 1.63, 27, 2022.

[14] R. Brignall and D. Cocks. A Framework for Minimal Hereditary Classes of Graphs of Unbounded Clique-Width. *SIAM J. Discrete Math.*, 37(4):2558–2584, 2023.

[15] R. Brignall, N. Korpelainen, and V. Vatter. Linear clique-width for hereditary classes of cographs. *J. Graph Theory*, 84(4):501–511, 2017.

[16] D. Cocks. t-sails and sparse hereditary classes of unbounded tree-width. *European J. Combin.*, 122:Paper No. 104005, 2024.

[17] A. Collins, J. Foniok, N. Korpelainen, V. V. Lozin, and V. Zamaraev. Infinitely many minimal classes of graphs of unbounded clique-width. *Discrete Appl. Math.*, 248:145–152, 2018.

[18] B. Courcelle. The monadic second-order logic of graphs. I. Recognizable sets of finite graphs. *Inform. and Comput.*, 85(1):12–75, 1990.

[19] B. Courcelle, J. Engelfriet, and G. Rozenberg. Handle-rewriting hypergraph grammars. *J. Comput. System Sci.*, 46(2):218–270, 1993.

[20] B. Courcelle, J. A. Makowsky, and U. Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory Comput. Syst.*, 33(2):125–150, 2000.

[21] B. Courcelle and S. Olariu. Upper bounds to the clique width of graphs. *Discrete Appl. Math.*, 101(1-3):77–114, 2000.

[22] K. K. Dabrowski, S. Huang, and D. Paulusma. Bounding clique-width via perfect graphs. *J. Comput. System Sci.*, 104:202–215, 2019.

[23] K. K. Dabrowski, M. Johnson, and D. Paulusma. Clique-width for hereditary graph classes. In *Surveys in combinatorics 2019*, volume 456 of *London Math. Soc. Lecture Note Ser.*, pages 1–56. Cambridge Univ. Press, Cambridge, 2019.

[24] K. K. Dabrowski, V. V. Lozin, and D. Paulusma. Clique-width and well-quasi-ordering of triangle-free graph classes. *J. Comput. System Sci.*, 108:64–91, 2020.

[25] K. K. Dabrowski, T. Masařík, J. Novotná, D. Paulusma, and P. Rzążewski. Clique-width: harnessing the power of atoms. *J. Graph Theory*, 104(4):769–810, 2023.

[26] K. K. Dabrowski and D. Paulusma. Clique-width of graph classes defined by two forbidden induced subgraphs. *Comput. J.*, 59(5):650–666, 2016.

[27] J. Daligault, M. Rao, and S. Thomassé. Well-quasi-order of relabel functions. *Order*, 27(3):301–315, 2010.

[28] T. Davies. Problem session 4, Treewidth of hereditary classes, p66, Oberwolfach technical report DOI:10.4171/OWR/2022/1.

[29] A. Dawar and A. Sankaran. MSO undecidability for hereditary classes of unbounded clique width. In *30th EACSL Annual Conference on Computer Science Logic*, volume 216 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 17, 17. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022.

[30] R. Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, Berlin, fifth edition, 2017.

[31] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[32] R. Fraïssé. Sur l'extension aux relations de quelques propriétés des ordres. *Ann. Sci. Ecole Norm. Sup. (3)*, 71:363–388, 1954.

[33] J. Geelen, O.-J. Kwon, R. McCarty, and P. Wollan. The grid theorem for vertex-minors. *Journal of Combinatorial Theory, Series B*, 2020.

[34] M. C. Golumbic and U. Rotics. On the clique-width of some perfect graph classes. *Internat. J. Found. Comput. Sci.*, 11(3):423–443, 2000. Selected papers from the Workshop on Theoretical Aspects of Computer Science (WG 99), Part 1 (Ascona).

[35] F. Gurski. The behavior of clique-width under graph operations and graph transformations. *Theory of Computing Systems*, pages 1–31, 2016.

[36] F. Gurski and E. Wanke. The tree-width of clique-width bounded graphs without $K_{n,n}$. In *Graph-theoretic concepts in computer science (Konstanz, 2000)*, volume 1928 of *Lecture Notes in Comput. Sci.*, pages 196–205. Springer, Berlin, 2000.

[37] F. Gurski and E. Wanke. On the relationship between NLC-width and linear NLC-width. *Theoret. Comput. Sci.*, 347(1-2):76–89, 2005.

[38] F. Gurski and E. Wanke. Line graphs of bounded clique-width. *Discrete Math.*, 307(22):2734–2754, 2007.

[39] R. Hickingbotham, F. Illingworth, B. Mohar, and D. R. Wood. Treewidth, circle graphs, and circular drawings. *SIAM J. Discrete Math.*, 38(1):965–987, 2024.

[40] W. Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[41] T. Korhonen. Grid induced minor theorem for graphs of small degree. *J. Combin. Theory Ser. B*, 160:206–214, 2023.

[42] N. Korpelainen. A new graph construction of unbounded clique-width. In *TCDM 2016—1st IMA Conference on Theoretical and Computational Discrete Mathematics, University of Derby*, volume 56 of *Electron. Notes Discrete Math.*, pages 31–36. Elsevier Sci. B. V., Amsterdam, 2016.

[43] V. V. Lozin. Boundary classes of planar graphs. *Combin. Probab. Comput.*, 17(2):287–295, 2008.

[44] V. V. Lozin. Minimal classes of graphs of unbounded clique-width. *Ann. Comb.*, 15(4):707–722, 2011.

[45] V. V. Lozin and M. Milanič. Critical properties of graphs of bounded clique-width. *Discrete Math.*, 313(9):1035–1044, 2013.

[46] V. V. Lozin and D. Rautenbach. Chordal bipartite graphs of bounded tree- and clique-width. *Discrete Math.*, 283(1-3):151–158, 2004.

[47] V. V. Lozin and D. Rautenbach. On the band-, tree-, and clique-width of graphs with bounded vertex degree. *SIAM J. Discrete Math.*, 18(1):195–206, 2004.

[48] V. V. Lozin and I. Razgon. Tree-width dichotomy. *European J. Combin.*, 103:Paper No. 103517, 8, 2022.

[49] V. V. Lozin, I. Razgon, and V. Zamaraev. Well-quasi-ordering versus clique-width. *J. Combin. Theory Ser. B*, 130:1–18, 2018.

[50] J. A. Makowsky and U. Rotics. On the clique-width of graphs with few P_4's. *Internat. J. Found. Comput. Sci.*, 10(3):329–348, 1999.

[51] R. McCarty. *Local Structure for Vertex-Minors*. PhD thesis, University of Waterloo, 2021.

[52] C. St. J. A. Nash-Williams. Decomposition of finite graphs into forests. *J. London Math. Soc.*, 39:12, 1964.

[53] S.-I. Oum. Rank-width and vertex-minors. *J. Combin. Theory Ser. B*, 95(1):79–100, 2005.

[54] S.-I. Oum and P. Seymour. Approximating clique-width and branch-width. *J. Combin. Theory Ser. B*, 96(4):514–528, 2006.

[55] A. C. Pohoata. "Unavoidable induced subgraphs of large graphs" Senior thesis, Princeton University (2014).

[56] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[57] N. Robertson and P. D. Seymour. Graph minors. V. Excluding a planar graph. *J. Combin. Theory Ser. B*, 41(1):92–114, 1986.

[58] N. Robertson and P. D. Seymour. Graph minors. XX. Wagner's conjecture. *J. Combin. Theory Ser. B*, 92(2):325–357, 2004.

[59] N. L. D. Sintiari and N. Trotignon. (Theta, triangle)-free and (even hole, K_4)-free graphs—part 1: Layered wheels. *J. Graph Theory*, 97(4):475–509, 2021.

[60] K. Wagner. Über eine Eigenschaft der ebenen Komplexe. *Math. Ann.*, 114(1):570–590, 1937.

[61] D. Weißauer. On the block number of graphs. *SIAM J. Discrete Math.*, 33(1):346–357, 2019.

[62] Z. X. Wen and Z. Y. Wen. Local isomorphisms of invertible substitutions. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(4):299–304, 1994.

[63] E. Zeckendorf. Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas. *Bull. Soc. Roy. Sci. Liège*, 41:179–182, 1972.