Abstract

The Gyárfás-Sumner conjecture \cite{29,42} states that for every tree $T$, there is a function $f$ such that for every graph $G$ with no induced subgraph isomorphic to $T$, the chromatic number of $G$ is at most $f(\omega(G))$, where $\omega(G)$ is its clique number. We prove this when $T$ is a tree formed by joining two disjoint paths by an edge.

A class $\mathcal{C}$ of graphs has the EH-property if there is a $\delta > 0$ such that every $G \in \mathcal{C}$ has a clique or stable set of size at least $|V(G)|^\delta$. The Erdős-Hajnal conjecture \cite{21,22} states that for every graph $H$, the class of $H$-free graphs has the EH-property. One approach for proving this is showing that there exists an $\varepsilon > 0$ such that every $G \in \mathcal{C}$ contains two sets $A, B$ with $|A|, |B| \geq \varepsilon|V(G)|$ and such that either no edges or all edges between the sets are present in $G$. We prove a conjecture of Liebenau and Pilipczuk \cite{33}, that for every tree $T$, the class of graphs containing neither $T$ nor its complement as an induced subgraph has this property, and thus has the EH-property. This generalizes several previous results \cite{4,7,33,34}. We consider variants obtained by requiring that $G$ is sparse, or $A, B$ have polynomial instead of linear size, or the density between $A$ and $B$ is bounded, or $G$ contains few copies of $H$. We prove a conjecture of Conlon, Fox and Sudakov \cite{18} for almost bipartite graphs. Our results imply an improved bound for the Erdős-Hajnal conjecture for excluding a five-cycle, the simplest open case.

The strong perfect graph theorem \cite{11} contains a decomposition theorem, and even though perfect graphs can be colored in polynomial time \cite{28}, no combinatorial algorithm for this is known. One obstacle for such an algorithm are “skew partitions”, which arise from induced subgraphs isomorphic to line graphs. Generalizations of line graphs, so-called orthogonal strip systems, yield particularly bad skew partitions. We prove that in a perfect graph, under mild assumptions, the number of pairwise orthogonal strip systems is bounded by twice the clique number.
Acknowledgements

First and foremost, I would like to thank my advisors, Maria Chudnovsky and Paul Seymour, for everything they have taught me, and for their mentorship and guidance. I am incredibly grateful for the countless hours we spent working together, which helped me become a better mathematician, researcher, and writer.

I am thankful to Jacob Fox, Aurélie Lagoutte, Anita Liebenau, Marcin Pilipczuk and Alex Scott for generously sharing their ideas.

My sincere thanks goes to Noga Alon, Maria Chudnovsky, and Paul Seymour for serving on my thesis committee, and to Alex Scott for being a thesis reader.

I am grateful to Princeton University and the Program in Applied and Computational Mathematics. I am especially thankful for Tina Dwyer, who has patiently guided me through the formal process for submitting and defending this dissertation.

I would also like to thank Gina Holland and Audrey Mainzer for all of their help with the administrative side of my studies.

I gratefully acknowledge the funding I received while at Princeton University. My research was partially supported by NSF grant DMS-1265563, NSF grant DMS-1550991. This material is based upon work supported in part by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-16-1-0404. During my final year of studies, I was supported by the Charlotte Elizabeth Procter fellowship, which has allowed me to devote more of my time to research.

I am thankful to my parents, Kathrin and Reinhard, and to my sister Nora, for their love and support. And finally, I would like to thank my husband Logan, for everything, but mostly for always being there for me.
# Contents

Abstract ................................................................. iii  
Acknowledgements ......................................................... iv  

1 Introduction .......................................................... 1  
1.1 Prior publication and joint work .................................. 5  
1.2 Definitions .......................................................... 6  
1.3 Tools ................................................................. 10  

2 Excluding an H-graph .................................................. 13  
2.1 Growing stars ........................................................ 15  
2.2 The 2-controlled case ............................................... 18  
2.3 The $r$-controlled case ............................................ 23  
2.4 The uncontrolled case .............................................. 39  

3 Complete or anticomplete pairs ..................................... 50  
3.1 The Erdős-Hajnal conjecture ...................................... 51  
3.1.1 Sparse symmetric version ...................................... 53  
3.2 Linear anticomplete pairs in sparse graphs ...................... 57  
3.2.1 Necessary conditions ........................................... 58  
3.2.2 Connectivity and growing paths .............................. 60  
3.2.3 Caterpillars ..................................................... 64  
3.2.4 Caterpillar subdivisions ....................................... 74
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.5</td>
<td>All trees</td>
<td>85</td>
</tr>
<tr>
<td>3.2.6</td>
<td>Directed graphs</td>
<td>112</td>
</tr>
<tr>
<td>3.2.7</td>
<td>Holes of fixed length mod (k)</td>
<td>117</td>
</tr>
<tr>
<td>3.3</td>
<td>Linear and polynomial anticomplete pairs in sparse graphs</td>
<td>119</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Holes</td>
<td>122</td>
</tr>
<tr>
<td>3.3.2</td>
<td>All trees</td>
<td>126</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Long exact subdivisions</td>
<td>129</td>
</tr>
<tr>
<td>3.4</td>
<td>Polynomial complete or anticomplete pairs</td>
<td>138</td>
</tr>
<tr>
<td>3.5</td>
<td>Sparse sets in sparse graphs</td>
<td>143</td>
</tr>
<tr>
<td>3.5.1</td>
<td>A game on a graph</td>
<td>144</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Many pairwise sparse pairs</td>
<td>148</td>
</tr>
<tr>
<td>3.5.3</td>
<td>Bipartite graphs and matchings</td>
<td>156</td>
</tr>
<tr>
<td>3.6</td>
<td>Fixed-size anticomplete pairs</td>
<td>170</td>
</tr>
<tr>
<td>3.7</td>
<td>Triangle-free graphs</td>
<td>174</td>
</tr>
<tr>
<td>4</td>
<td>Orthogonal strip systems in perfect graphs</td>
<td>179</td>
</tr>
<tr>
<td>4.1</td>
<td>Properties of orthogonal strip systems</td>
<td>184</td>
</tr>
<tr>
<td>4.2</td>
<td>Finding a clique</td>
<td>190</td>
</tr>
<tr>
<td>5</td>
<td>Conclusion</td>
<td>197</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>200</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In this thesis, we consider three problems related to graphs with forbidden induced subgraphs.

The first problem is the Gyárfás-Sumner conjecture, which was proposed independently by Gyárfás [29] and Sumner [42]. This conjecture states that for every tree $T$, there is a function $f$ such that for every $T$-free graph $G$, the chromatic number of $G$ is bounded by $f(\omega(G))$, where $\omega(G)$ is the clique number of $G$. This conjecture is open, although it has been proved for several classes of trees [14, 29–31, 39]. In Chapter 2, we prove the Gyárfás-Sumner conjecture for a class of trees called H-graphs, which are trees that are constructed by joining two disjoint paths by an edge. In all but one case, our techniques also work for trees obtained by joining the centers of two subdivided stars by an edge.

In Chapter 3, we consider a second problem, which is the Erdős-Hajnal conjecture. This conjecture, due to Erdős and Hajnal [21,22], states that if $H$ is a graph, then every graph $G$ that does not contain $H$ as an induced subgraph has a polynomial-size set of vertices that are either all pairwise adjacent, or all pairwise non-adjacent. Many attempts for proving this conjecture and its variants are based on finding two vertex sets with high or low edge density between them in an $H$-free graph [347].
We consider several variants for different densities and desired sizes of the two vertex sets.

In Section 3.1.1, we use a result of [38] to show that if we exclude a graph $H$ and its complement, then we may assume that $G$ is sparse, i.e. every vertex of $G$ is adjacent to at most a small fraction of the vertex set of $G$. In Section 3.2 we ask for which graphs $H$ it is true that every $H$-free sparse graph $G$ contains two linear-size sets $A, B$ with no edges between them. In this context, “linear” always refers to a linear function in terms of $|V(G)|$, the number of vertices in $G$. Liebenau and Pilipczuk [33] conjectured that this is true if and only if $H$ is a forest, which we prove in Section 3.2.5. In Section 3.2.3 we present a separate, simpler proof for a class of trees called caterpillars, which leads to the result of Section 3.2.4 that excluding all subdivisions of a graph $H$ from a sparse graph $G$ guarantees that $G$ contains two linear-size sets $A, B$ with no edges between them. We also consider this question in directed graphs in Section 3.2.6, and show that this property fails to hold even for very simple directed graphs, such as an in-directed or out-directed claw and a directed two-edge path. Finally, in Section 3.2.7 we prove that the property holds when excluding all holes with length congruent to $j$ modulo $k$ if and only if $j$ is even or $k$ is odd, where “only if” follows from a result of [25] showing that even sparse perfect graphs can fail to contain two linear-size sets $A, B$ with no edges between them.

In Section 3.3 we consider a relaxation of this problem: for which graphs $H$ does every sparse $H$-free graph contain two sets $A, B$ with no edges between them, where $A$ has linear size and $B$ has polynomial size? Conlon, Fox, and Sudakov [18] conjectured that this holds for all graphs $H$. We prove that it is true if $H$ is a hole or a long subdivision of any graph, and we also give a simpler proof that this holds for all trees. In this context, a “density version” of the question is also helpful, where instead we ask for a $k \in \mathbb{N}$ such that for all $c > 0$, there are two sets $A$ and $B$ in $G$ such that
there are at most $c|A||B|$ edges between $A$ and $B$, and $|A| \geq c^k|V(G)|$, and $B$ has linear size. We prove that for holes of length at least five, and for long subdivisions of any graph, this is true with $k = 1$.

Section 3.4 contains a further relaxation; now we ask for both $A$ and $B$ to have polynomial size, and we also allow the outcome in which all edges between $A$ and $B$ are present. This is true for every graph $H$ we exclude, and was proved by Erdős, Hajnal and Pach [23]. With similar methods, we strengthen this and prove that there is an $\varepsilon > 0$ depending on $H$ such that every sufficiently sparse $H$-free graph contains sets $A$ and $B$ with $|A||B| \geq |V(G)|^{1+\varepsilon}$. Again, we also consider a density version; now this means that for all $c > 0$, there is a $k \in \{0, \ldots, 2^{|V(H)|}\}$ such that $G$ contains two sets $A$, $B$ with of size at least a constant times $c^k|V(G)|$, and there are at most $c^{k+1}|A||B|$ or at least $(1 - c^{k+1})|A||B|$ edges between $A$ and $B$. We prove that this density version is true for every graph $H$ we exclude.

In Section 3.5 we improve the results from Section 3.4 by proving that for every graph $H$, if $G$ is $H$-free and sufficiently sparse, then $G$ contains sets $A$ and $B$ of polynomial size and with no edges between them, and in fact, there is an $\varepsilon > 0$ such that $|A||B| \geq |V(G)|^{1+\varepsilon}$. By using a “$k$-tuple game”, we also prove that for every $k$, every sufficiently sparse $H$-free graph contains sets $A_1, \ldots, A_k$, each of polynomial size, and with no edges between them. These results are strengthened by a density version in two ways; we prove that for every $H$ and $k$, there is a $t \in \mathbb{N}$ such that if $G$ is sufficiently sparse, then for all $c > 0$, either $G$ contains at least $c^t|V(G)||V(H)|$ distinct copies of $H$, or $k$ disjoint sets $A_1, \ldots, A_k$, each of size at least $c^t|V(G)|$, such that for all distinct $i, j \in \{1, \ldots, k\}$, the number of edges between $A_i$ and $A_j$ is at most $c|A_i||A_j|$. These results carry over to the non-sparse setting using a result of [27], which shows that for every graph $H$, there is a $\kappa > 0$ such that every graph $G$ contains either at least $\kappa|V(G)||V(H)|$ copies of $H$, or a linear-size induced subgraph $J$ such that $J$ or $J^c$ is sparse.
In the second half of Section 3.5 we use an even stronger version of the density result from Section 3.4 in which the number of edges between two sets is even smaller (the bound involves a different power of $c$) to derive a proof that the question in Section 3.3 can be answered in the affirmative when $H$ is a bipartite graph. We prove that in this case, there is a $k \in \mathbb{N}$ such that for all sufficiently sparse $G$, and for all $c > 0$, either $G$ contains at least a constant times $c^k |V(G)|$ copies of $H$, or there are two sets $A$ and $B$ in $G$ such that there are at most $c|A||B|$ edges between $A$ and $B$, and $|A| \geq c^k |V(G)|$, and $B$ has linear size. We prove a similar result, albeit with only one copy of $H$, for almost bipartite graphs, which are triangle-free graphs whose vertex set can be partitioned into a stable set and the vertex set of a matching. This implies that the question in Section 3.3 can be answered affirmatively for every graph $H$ that is a subdivision of a graph $H'$ in which every edge has been subdivided at least once.

In Section 3.6 we revisit the questions of Sections 3.3 and 3.4 and consider upper bounds on the polynomial size of the sets $A$ and $B$. We prove that the same polynomial does not work for all graphs $H$. By the result of Section 3.2.5 it follows that linear-size sets work for trees; here we give a simple proof that sets of size a constant times $\sqrt{|V(G)|}$ work for all trees if $G$ is triangle-free.

Finally, in Section 3.7 we summarize our results for excluding a triangle, which is the simplest and smallest graph for which we cannot prove the conjecture of [18]. In a triangle-free graph, the set of neighbors of any vertex is a stable set, and hence the assumption that $G$ is sparse is no longer necessary. We prove that every triangle-free graph $G$ contains two sets, one of linear size, one of size $\approx \frac{\log |V(G)|}{\log \log |V(G)|}$, with no edges between them. We also prove an upper bound; there are triangle-free graphs without a set of size a constant times $\sqrt{|V(G)|}$ and a set of linear size with no edges between them.
The third and last problem we consider is the problem of finding a polynomial-time combinatorial coloring algorithm for perfect graphs. A graph $G$ is perfect if for each of its induced subgraphs $H$, the clique number of $H$ equals its chromatic number. The celebrated Strong Perfect Graph Theorem due to Chudnovsky, Robertson, Seymour and Thomas [11] includes a structural description of this class of graphs, and Grötschel, Lovász and Schrijver [28] showed that an optimum coloring for a perfect graph can be found in polynomial time using the ellipsoid method. However, no polynomial-time combinatorial algorithm for this problem is known. In Chapter 4 we consider a problem related to coloring perfect graphs. In the structural description, one of the elements that has been an obstacle for polynomial-time algorithms are balanced skew partitions arising from line graphs [10, 17]. These balanced skew partitions are particularly problematic if the line graphs form so-called “orthogonal strip systems”. We prove that the number of pairwise orthogonal strip systems in a perfect graph is bounded by twice its clique number.

1.1 Prior publication and joint work

Chapter 1 contains standard graph theory notation and well-known results. In the remainder of this Chapter, we introduce definitions (Section 1.2) and tools (Section 1.3) that will be used throughout this thesis. Chapter 2 is based on joint work with Maria Chudnovsky, Alex Scott, and Paul Seymour. These results have not been submitted for publication.

Sections 3.1, 3.1.1, and the beginning of Section 3.2 contain a review of previous results. In Section 3.2.1 we review a construction of [20] and give a proof for completeness. In Section 3.2.2 we give an alternate proof of a result of [4] and some further tools based on joint work with Maria Chudnovsky, Alex Scott and Paul Sey-
mour. Section 3.2.3 is based on joint work with Anita Liebenau, Marcin Pilipczuk, and Paul Seymour, and the results have been submitted for publication [34]. Sections 3.2.4, 3.2.5, 3.2.6 and 3.2.7 are based on joint work with Maria Chudnovsky, Alex Scott and Paul Seymour. The result of Section 3.2.4 has been submitted for publication [16]. Sections 3.3, 3.4, 3.5, 3.6 and 3.7 are based on joint work with Maria Chudnovsky, Jacob Fox, Alex Scott, and Paul Seymour. The result of Section 3.3.1 has been submitted for publication [9].

Chapter 4 is based on joined work with Maria Chudnovsky, Aurélie Lagoutte, and Paul Seymour, which has not been submitted for publication. Chapter 5 contains concluding remarks.

1.2 Definitions

A graph $G$ consists of a (finite) set of vertices $V(G)$, and a set $E(G) \subseteq \binom{V(G)}{2}$ of edges. Edges are denoted as $uv$, where $u, v \in V(G)$, and if $uv \in E(G)$, then $u$ and $v$ are called adjacent in $G$. All edges are undirected unless stated otherwise, which means that $uv$ denotes the same edge as $vu$. An induced subgraph of a graph $G$ is a graph $H$ with vertex set $V(H) \subseteq V(G)$ and such that $u, v \in V(H)$ are adjacent in $H$ if and only if $u$ and $v$ are adjacent in $G$. For a graph $H$ and a graph $G$, we say that $G$ contains $H$ (as an induced subgraph) if there is an induced subgraph of $G$ isomorphic to $H$; otherwise we say that $G$ is $H$-free. For a set $\mathcal{F}$ of graphs, we say that $G$ is $\mathcal{F}$-free if $G$ is $F$-free for all $F \in \mathcal{F}$.

Let $G$ be a graph and $X \subseteq V(G)$. We use $G \setminus X$ to denote the graph that arises from $G$ by deleting every vertex in $X$, i.e. the graph with vertex set $V(G) \setminus X$ and edge set $\{e = uv \in E(G) : u, v \in V(G) \setminus X\}$. If $X = \{x\}$, we also write $G \setminus x$ for $G \setminus \{x\}$. Moreover, the graph $G$ restricted to $X$, denoted $G|X$, is defined as $G \setminus (V(G) \setminus X)$. For $X \subseteq V(G)$, we write $E(X)$ for $E(G|X)$; for disjoint $X, Y \subseteq V(G)$, we write
\[ E(X,Y) = \{ e \in E(G) : e = uv, u \in X, v \in Y \} \] for the sets of edges between \( X \) and \( Y \).

For an edge \( e \in E(G) \), we let \( G \setminus e \) denote the graph with vertex set \( V(G) \) and edge set \( E(G) \setminus \{ e \} \).

Let \( G \) be a graph. For disjoint sets \( X, Y \subseteq V(G) \), we say that \( X \) is complete to \( Y \) if \( xy \in E(G) \) for every \( x \in X \) and \( y \in Y \). We say that \( X \) is anticomplete to \( Y \) if \( xy \not\in E(G) \) for every \( x \in X \) and \( y \in Y \). For \( v \in V(G) \) and \( X \subseteq V(G) \) with \( v \not\in X \), we say that \( v \) is complete to \( X \) (and \( X \) is complete to \( v \)) if \( \{ v \} \) is complete to \( X \), and \( v \) is anticomplete to \( X \) (and \( X \) is anticomplete to \( v \)) if \( \{ v \} \) is anticomplete to \( X \).

Let \( G \) be a graph. For disjoint sets \( X, Y \subseteq V(G) \), we say that \( X \) is complete to \( Y \) if \( xy \in E(G) \) for every \( x \in X \) and \( y \in Y \). We say that \( X \) is anticomplete to \( Y \) if \( xy \not\in E(G) \) for every \( x \in X \) and \( y \in Y \). For \( v \in V(G) \) and \( X \subseteq V(G) \) with \( v \not\in X \), we say that \( v \) is complete to \( X \) (and \( X \) is complete to \( v \)) if \( \{ v \} \) is complete to \( X \), and \( v \) is anticomplete to \( X \) (and \( X \) is anticomplete to \( v \)) if \( \{ v \} \) is anticomplete to \( X \).

A path of length \( k \) is a graph \( P \) with \( k+1 \) vertices \( v_1, \ldots, v_{k+1} \) and with \( E(P) = \{ v_i v_{i+1} : i \in \{1, \ldots, k\} \} \). It is an odd path if \( k \) is odd, and an even path, otherwise. We also write \( P = v_1-v_2-\ldots-v_k-v_{k+1} \) to denote a path with vertices \( v_1, v_2, \ldots, v_k, v_{k+1} \) in order; \( v_1 \) and \( v_{k+1} \) are its ends, and \( P \) is a \( v_1-v_{k+1} \)-path. For a graph \( G \), if \( v \) is an end of \( P \), and \( P \) is an induced subgraph of \( G \), we say that \( P \) is a path in \( G \) starting at \( v \). Moreover, if \( u, v \in V(P) \), we write \( uPv \) for the \( u-v \)-path contained in \( P \). We denote the path of length \( k \) as \( P_{k+1} \). For a \( u-v \)-path \( P \), we say that \( V(P) \setminus \{ u, v \} \) is the interior of \( P \). An antipath is a graph that is the complement of a path. Its length is defined as the length of the path that is its complement.

A cycle of length \( k \) is a graph with \( k \) vertices \( v_1, \ldots, v_k \) and with \( E(G) = \{ v_i v_{i+1} : i \in \{1, \ldots, k-1\} \} \cup \{ v_1 v_k \} \). We denote the cycle of length \( k \) as \( C_k \). A cycle is called a hole if its length is at least four. A graph \( H \) is an antihole if \( H^c \) is a hole. A hole is odd if its length is odd, and an antihole is odd if the length of its complement is odd. A cycle of length three is called a triangle.
For a graph $G$ and $u,v \in V(G)$, we let $d_G(u,v)$ (or $d(u,v)$ if $G$ is clear from context) denote the distance from $u$ to $v$, i.e. the length of a shortest $u$-$v$-path in $G$ (where we define $d_G(u,v) = \infty$ if no such path exists). We say that $G$ is connected if $d_G(u,v) \neq \infty$ for all $u,v \in V(G)$. An induced subgraph $H$ of $G$ is a connected component of $G$ if $H$ is connected and $G|(V(H) \cup \{v\})$ is not connected for all $v \in V(G) \setminus V(H)$. A graph $G$ is anticonnected if $G^c$ is connected. An induced subgraph $H$ of $G$ is an anticomponent of $G$ if $H^c$ is a connected component of $G^c$.

For two graphs $G$ and $G'$, their disjoint union $G \cup G'$ is the graph with vertex set $V(G) \cup V(G')$ and edge set $E(G) \cup E(G')$. Their join is the graph $(G^c \cup G'^c)^c$, i.e. the graph that arises from the disjoint union of $G$ and $G'$ by adding all edges between $V(G)$ and $V(G')$.

For a graph $G$, a graph $H$ is a subdivision of $G$ if $V(G) \subseteq V(H)$ and

- for every edge $e = uv \in E(G)$, there is a path $P_e$ in $H$ such that $P_e$ is an induced $u$-$v$-path;

- for every edge $e \in E(G)$, the interior of $P_e$ is disjoint from $V(G)$ and anticomplete to $V(G) \setminus \{u,v\}$ in $H$;

- for all $e,f \in E(G)$ with $e \neq f$, the interior of $P_e$ is disjoint from the interior of $P_f$ and anticomplete to the interior of $P_f$ in $H$;

- for all $v \in V(H) \setminus V(G)$, there exists an edge $e \in E(G)$ and a path $P_e$ such that $v$ is in the interior of $P_e$; and

- for all $f \in E(H)$, there exists an edge $e \in E(G)$ and a path $P_e$ such that $f \in E(P_e)$.

For every edge $e \in E(G)$, $P_e$ is a branch of $H$.

For a vertex $v \in V(G)$, we write $N_G(v)$ (or $N(v)$ if $G$ is clear from context) for the set of neighbors of $v$, i.e. the set of $w \in V(G)$ such that $vw \in E(G)$. In
other words, $N(v)$ is the set of vertices in $G$ with distance 1 from $v$ in $G$. We use $d_G(v)$ (or $d(v)$ if $G$ is clear from context) to denote the degree of $v$ and set $d(v) = |N(v)|.$ For $r \geq 0$, we define the $r$-th neighborhood of $v$ as $N^r(v) = N^r_G(v) = \{w \in V(G) : d(v, w) = r\}$. Moreover, the closed $r$-th neighborhood of $v$ is defined as $N^r[v] = N^r_G[v] = \{v\} \cup N(v) \cup \cdots \cup N^r(v) = \{w \in V(G) : d(v, w) \leq r\}$. We let $N[v] = N^1[v]$. For a graph $G$ and $A \subseteq V(G)$, we say that $v \in A$ is an $r$-center of $A$ if $A = N^r_{G \mid A}[v]$, i.e. every vertex of $A$ has distance at most $r$ from $v$ in $G \mid A$.

For a graph $G$ and $X \subseteq V(G)$, we write $N(X) = (\bigcup_{x \in X} N(x)) \setminus X$ for the neighbors of $X$. For $r \geq 0$, we define the $r$-th neighborhood of $X$ as $N^r(X) = \{w \in V(G) : r = \min_{x \in X} d(w, x)\}$. Moreover, the closed $r$-neighborhood of $X$ is defined as $N^r[X] = \{w \in V(G) : \min_{x \in X} d(w, x) \leq r\}$. It follows that if $X = x$, then $N(x) = N(X)$ and $N^r(x) = N^r(X)$ for all $r \geq 0$.

A graph is a forest if it does not contain a cycle. A forest is a tree if it is connected. A vertex of degree one in a tree is a leaf of the tree. For $d \geq 2$ and $h \geq 1$, we denote by $T_{d,h}$ the $d$-ary tree of height $h$, defined as a tree $T$ that has a root, i.e. a vertex $v \in V(T)$ such that $d(v) = d$, and in which $d(w) \in \{1, d + 1\}$ for all $w \in V(T) \setminus \{v\}$, and such that $d(v, w) = h$ for every leaf $w$ of $T$.

For a graph $G$, a set $S \subseteq V(G)$ is stable if $uv \not\in E(G)$ for all $u, v \in S$, and $S$ is a clique if $uv \in E(G)$ for all distinct $u, v \in S$. The stability number $\alpha(G)$ is defined as the size of a largest stable set in $G$, and the clique number $\omega(G)$ is the size of a largest clique in $G$. If $V(G)$ is a clique, we say that $G$ is a complete graph and write $G = K_{|V(G)|}$. For $s, t \geq 1$, we let $K_{s,t}$ be the complete bipartite graph, i.e. a graph with vertex set $A \cup B$ such that $A$ is a stable set of size $s$, $B$ is a stable set of size $t$, and $A$ is complete to $B$.

For a set $X$, a partition $M$ of $X$ is a set of disjoint subsets of $X$ such that $\bigcup_{S \in M} S = X$. A graph $G$ is bipartite if there is a partition $\{A, B\}$ of $V(G)$ such that
\[ E(G) = E(A, B) \]. Such a partition of a bipartite graph is called a bipartition. In particular, this implies that \( A \) and \( B \) are stable sets in \( G \).

Let \( G \) be a graph. A function \( f : V(G) \to \mathbb{N} \) is a (proper) coloring of \( G \) if \( f(u) \neq f(v) \) for all \( uv \in E(G) \). The graph \( G \) is \( k \)-colorable if there is a proper coloring \( f \) of \( G \) with \( f(V(G)) \subseteq \{1, \ldots, k\} \). The smallest number \( k \) such that \( G \) is \( k \)-colorable is the chromatic number \( \chi(G) \). For a graph \( G \) and \( X \subseteq V(G) \), we often write \( \chi(X) \) to denote \( \chi(G|X) \).

A class \( \mathcal{C} \) of graphs is a set of graphs. It is hereditary if it is closed under taking induced subgraphs, i.e. for every graph \( G \in \mathcal{C} \) and \( X \subseteq V(G) \), we have that \( G \setminus X \in \mathcal{C} \). By definition, for every set \( \mathcal{F} \) of graphs, the class of \( \mathcal{F} \)-free graphs is hereditary.

A directed graph (digraph) \( D \) consists of a finite set \( V(D) \) of vertices, and a set \( E(D) \subseteq (V(D) \times V(D)) \setminus \{(v, v) : v \in V(D)\} \) of edges. The underlying undirected graph of \( D \) is the graph \( G \) with \( V(G) = V(D) \) and \( E(G) = \{uv : (u, v) \in E(D)\} \). For a vertex \( v \in V(D) \), we let \( N_D^+(v) = \{u \in V(D) : (u, v) \in E(D)\} \) denote the in-neighbors of \( v \) in \( D \), and we let \( N_D^-(v) = \{u \in V(D) : (v, u) \in E(D)\} \) denote the out-neighbors of \( v \) in \( D \); we write \( N^+(v) \) and \( N^-(v) \) if \( D \) is clear from context. We write \( d^+(v) \) for the in-degree of \( v \) and let \( d^+(v) = |N^+(v)| \); and we write \( d^-(v) \) for the out-degree of \( v \) and let \( d^-(v) = |N^-(v)| \).

### 1.3 Tools

This section contains a few useful and well-known results that will be used throughout the thesis.

**Theorem 1.3.1** (Ramsey [37]). For all \( k, r \) and for all \( c_1, \ldots, c_k \in \mathbb{N} \), there is a number \( R_r(c_1, \ldots, c_k) \) such that given

- \( n \geq R_r(c_1, \ldots, c_k) \) and a set \( V \) with \( |V| = n \); and

- a function \( f \) that assigns a number in \( \{1, \ldots, k\} \) to every subset of \( V \) of size \( r \),
there exists an \( i \in \{1, \ldots, k\} \) and \( S \subseteq V \) such that \(|S| \geq c_i\) and \( f(T) = i \) for every \( T \subseteq S \) with \(|T| = r\).

In particular, for all \( a, b \in \mathbb{N} \), there is a number \( R(a, b) = R_2(a, b) \) such that for every graph \( G \) with at least \( R(a, b) \) vertices, \( G \) either has a stable set of size \( a \) or a clique of size \( b \).

The number \( R(c_1, \ldots, c_k) = R_2(c_1, \ldots, c_k) \) is called the Ramsey number of \( c_1, \ldots, c_k \).

The following three lemmas are very simple, and we give short proofs for completeness. The first lemma shows that for a graph \( G \) and \( A \subseteq V(G) \), if \( A \) has an \( r \)-center, then there is a vertex \( a \in A \) such that \( A \setminus \{a\} \) still has an \( r \)-center.

**Lemma 1.3.2.** Let \( G \) be a graph, \( A \subseteq V(G) \), and let \( v \) be an \( r \)-center of \( A \). If \(|A| \geq 2\), then there exists a vertex \( a \in A \) such that \( v \) is an \( r \)-center of \( A \setminus \{a\} \).

**Proof.** Let \( G \) be a graph, \( A \subseteq V(G) \), and let \( v \) be an \( r \)-center of \( A \). Without loss of generality, we may assume that \( V(G) = A \). For \( i \in \{0, \ldots, r\} \), we let \( L_i = N^i(v) \).

By our assumption that \( v \) is an \( r \)-center of \( A \), it follows that \( A = L_0 \cup \cdots \cup L_r \). Let \( k \in \{0, \ldots, r\} \) be maximum such that \( L_k \neq \emptyset \). If \( k = 0 \), then \( A = \{v\} \) and the result holds. Thus we may assume that \( k \neq 0 \). Let \( a \in L_k \). Since every vertex in \( L_j \) has a neighbor in \( L_{j-1} \) for all \( j \in \{1, \ldots, k\} \), it follows that for every vertex \( b \in L_j \), every shortest \( b-v \)-path has its interior in \( L_1 \cup \cdots \cup L_j \). Therefore, \( a \) is not in the interior of a shortest \( w-v \)-path for all \( w \in V(G) \setminus \{a\} \), and so for all \( w \in V(G) \setminus \{a\} \), it follows that \( d_{G \setminus \{a\}}(w, v) = d_G(w, v) \). This implies that \( v \) is an \( r \)-center of \( A \setminus \{a\} \).

The next lemma is similar to Lemma 1.3.2 and shows that every connected graph \( G \) has a vertex whose deletion keeps \( G \) connected.

**Lemma 1.3.3.** Let \( G \) be a connected graph with \(|V(G)| = n\). Then there exists a labeling \( \{v_1, \ldots, v_n\} = V(G) \) such that for all \( k \in \{1, \ldots, n\} \), \( G \setminus \{v_1, \ldots, v_k\} \) is connected.
Proof. We prove this by induction on $|V(G)| = n$. If $|V(G)| = 1$, the result holds. Now let $|V(G)| > 1$, and let $v \in V(G)$. Since $G$ is connected, it follows that $v$ is an $n$-center for $V(G)$. By Lemma 1.3.2 it follows that there is a vertex $a \in V(G)$ such that $v$ is an $n$-center for $V(G) \setminus \{a\}$, and so $G \setminus \{a\}$ is connected. By induction, there exists an labeling $\{v_1, \ldots, v_{n-1}\}$ of $V(G) \setminus \{a\}$ such that the lemma holds for $G \setminus \{a\}$. Now let $v_n = a$. Since $G \setminus \{a\}$ is connected, the result of the lemma follows.

The following lemma shows that if $A$ has an $r$-center, then every vertex in $A$ is a $2r$-center for $A$.

Lemma 1.3.4. Let $G$ be a graph, $A \subseteq V(G)$, and let $v$ be an $r$-center for $A$. Let $w \in A$. Then $w$ is a $2r$-center for $A$.

Proof. Let $G$ be a graph, $A \subseteq V(G)$, and let $v$ be an $r$-center for $A$; let $w \in A$, and let $u \in A$. By the triangle inequality, we have that $d_{G|A}(u, w) \leq d_{G|A}(u, v) + d_{G|A}(v, w) \leq r + r = 2r$. This implies the result.
Chapter 2

Excluding an H-graph

A class $\mathcal{C}$ of graphs is $\chi$-bounded if there is a function $f : \mathbb{N} \to \mathbb{N}$ such that for every $G \in \mathcal{C}$, we have that $\chi(G) \leq f(\omega(G))$.

The results in this section relate to the following conjecture, which, if proved, would answer the question for which graphs $H$ the class of $H$-free graphs is $\chi$-bounded.

Conjecture 2.0.1 (Gyárfás [29], Sumner [42]). For every graph $H$, the class of $H$-free graphs is $\chi$-bounded if and only if $H$ is a forest.

A classical result of Erdős [20] is the following:

Theorem 2.0.2 (Erdős). For every $g, c \geq 3$, there is a graph $G$ with $\chi(G) > c$ and such that $G$ does not contain a cycle of length at most $g$.

This result implies that the “only if” direction of Conjecture 2.0.1 holds, as observed by Gyárfás [29]: A graph $G$ as in Theorem 2.0.2 has no cycle of length at most three, and therefore it satisfies that $\omega(G) \leq 2$; and moreover, it does not contain any graph $H$ with a cycle of length at most $g$ as an induced subgraph. Now let $H$ be a graph for which the class of $H$-free graphs is $\chi$-bounded, say $\chi(G) \leq f(\omega(G))$ for every $H$-free graph $G$. We now apply Theorem 2.0.2 with $g = |V(H)|$ and $c = f(2)$ and obtain a triangle-free graph $G$ with $\chi(G) > f(2)$. It follows that $\chi(G) > f(\omega(G))$, ...
and thus $G$ contains $H$ as an induced subgraph. Since $G$ contains no cycle of length at most $|V(H)|$, it follows that $H$ contains no cycles, and therefore, $H$ is a forest.

Therefore, proving Conjecture 2.0.1 can be reduced to proving that the condition that $H$ is a forest is sufficient to guarantee that the class of $H$-free graphs is $\chi$-bounded. Since every forest is an induced subgraph of a tree, it would be sufficient to prove Conjecture 2.0.1 for all trees. The conjecture is open, but it has been proved for many classes of trees, including paths [29], trees that arise from stars by subdivision [39], trees with a 2-center [31], some trees with a 3-center [30], as well as a few other classes of trees [14]. All these trees either have a 3-center or share a common feature: There is at most one vertex in them with more than two neighbors of degree more than one. In the following, we will prove the conjecture for one further class of trees that contains trees that do not have this feature, and that do not have a 3-center.

A tree $T$ is a subdivided star if $T$ has at most one vertex of degree more than two. A vertex of maximum degree in a subdivided star $T$ is called its center, and if $v$ is a center of $T$, we say that $T$ is centered at $v$. A subdivided star $T$ is an $l$-star if $T\setminus\{v\}$ consists of $l$ paths of length $l - 1$, where $v$ is the center of $T$. Note that an $l$-star has $l^2 + 1$ vertices. A tree $T$ is a double star if $T$ has at most two vertices of degree more than two, and if $T$ has two such vertices, then they are adjacent. In particular, every double star consists of the disjoint union of two subdivided stars $T_1, T_2$ along with an edge $uv$ joining a center $u$ of $T_1$ to a center $v$ of $T_2$. For every double star $T$, there is an $l \in \mathbb{N}$ such that $T$ is contained in the double star obtained by joining the centers of two $l$-stars by an edge.

A double star is called an $H$-graph if its maximum degree is at most three. Our main result is the following:

**Theorem 2.0.3.** For every $H$-graph $T$, the class of $T$-free graphs is $\chi$-bounded.
In the following, we fix a double star $T$ and $l \in \mathbb{N}$ such that $T$ is contained in the double star obtained by joining the centers of two $l$-stars by an edge; in Section 2.4 we further assume that $T$ is an $H$-graph.

For $r \in \mathbb{N}$ and a non-decreasing function $\varphi : \mathbb{N} \to \mathbb{N}$, we say that $G$ is $(r, \varphi)$-controlled if for every induced subgraph $H$ of $G$, there exists a vertex $v \in V(H)$ such that $\chi(H) \leq \varphi(\chi(N^r_H[v]))$.

The proof proceeds by induction on $\kappa = \omega(G)$. We fix $\kappa \geq 2$, and by induction, we may assume that there exists a constant $\tau$ such that $\chi(G) \leq \tau$ for every $T$-free graph $G$ with $\omega(G) \leq \kappa - 1$. In particular, this implies that $\chi(N^1(v)) \leq \tau$ for all $v \in V(G)$ and all graphs $G$ with $\omega(G) \leq \kappa$. In Section 2.1 we prove some useful lemmas that will be used throughout the proof. In Section 2.2 we show that for every function $\varphi$, the class of $(2, \varphi)$-controlled $T$-free graphs is $\chi$-bounded. In Section 2.3 we show that for every function $\varphi$, and every $r \geq 3$, the class of $(r, \varphi)$-controlled $T$-free graphs is $\chi$-bounded. Finally, in Section 2.4 we will show that there is an $r$ and a function $\varphi$ such that the class of $T$-free graphs that are not $(r, \varphi)$-controlled is $\chi$-bounded.

### 2.1 Growing stars

We begin with two simple lemmas that give sufficient conditions for the presence of certain induced subgraphs. The first lemma is well-known, and the proof closely follows the proof of Theorem 2.4 in [29]:

**Lemma 2.1.1.** Let $G$ be a connected graph and $v \in V(G)$. Let $\tau > 0$ and suppose that $\chi(N[w]) \leq \tau$ for all $w \in V(G)$. Then there exists an induced path $P$ in $G$ starting at $v$, and $P$ has at least $1 + \lfloor \chi(G)/\tau \rfloor$ vertices.

**Proof.** Let $k = \lfloor \chi(G)/\tau \rfloor$; we proceed by induction on $k$. If $k = 0$, then the result follows by choosing $V(P) = \{v\}$. Now suppose that $k \geq 1$. Then $\chi(G \setminus N[v]) \geq \chi(G) - \tau > 0$. Let $C$ be a component of $G \setminus N[v]$ with $\chi(C) = \chi(G \setminus N[v])$. It follows
that $V(C) \neq \emptyset$. Since $G$ is connected, it follows that there exists a vertex $v' \in N(v)$ such that $v'$ has a neighbor in $V(C)$. By induction, it follows that there exists an induced path $P'$ in $G|\{\{v'\} \cup V(C)\}$ starting in $v'$ with at least $1 + \lceil \chi(C)/\tau \rceil \geq 1 + \lceil (\chi(G) - \tau)/\tau \rceil = k$ vertices. Since $N(v) \cap \{\{v'\} \cup V(C)\} = \{v'\}$, it follows that $N(v) \cap V(P') = \{v'\}$; thus $v-v'P'$ is an induced path and has length at least $k$. This concludes the proof.

The next lemma is a variant of Theorem 6.2 in [14].

**Lemma 2.1.2.** Let $l > 0$ be fixed. Let $G$ be a graph and $r, \tau > 1$ such that $\chi(N^r[v]) \leq \tau$ for all $v \in V(G)$. Suppose that there exists a vertex $v \in V(G)$ such that $\chi(N^{r+1}(v)) > (l+1)^3\tau^2$. Then $G$ contains an $l$-star with center $v$ as an induced subgraph.

**Proof.** Let $v$ be as in the statement of the lemma. Let $\tau' = (l+1)^3\tau$. Since $\chi(N(v)) \leq \tau$, it follows $N(v)$ can be partitioned into $\tau$ stable sets $B_1, \ldots, B_\tau$. Therefore, there exists an $i \in \{1, \ldots, \tau\}$ such that $\chi(N^{r+1}(v) \cap N^r(B_i)) \geq \chi(N^{r+1}(v))/\tau = \tau'$; we let $B = B_i$ and let $C = N^{r+1}(v) \cap N^r(B)$.

We choose $A_1 \subseteq B$ minimal with respect to inclusion such that $\chi(N^r(A_1) \cap C) > l^2\tau$. We let $C_1 = N^r(A_1) \cap C$. For $2 \leq i \leq l$, we choose $A_i \subseteq (B \setminus (A_1 \cup \cdots \cup A_{i-1}))$ minimal with respect to inclusion such that

$$\chi(N^r(A_i) \cap (C \setminus (C_1 \cdots \cup C_{i-1}))) > l^2\tau.$$

We let $C_i = N^r(A_i) \cap (C \setminus (C_1 \cdots \cup C_{i-1}))$. Since $\chi(N^r(w)) \leq \tau$ for all $w \in V(G)$, it follows that $\chi(C_1 \cup \cdots \cup C_i) \leq i(l^2 + 1)\tau$ for all $i \in \{0, \ldots, l\}$, and hence

$$\chi(N^r(N(v) \setminus (A_1 \cup \cdots \cup A_l)) \cap (C \setminus (C_1 \cup \cdots \cup C_l))) \geq \tau' - i(l^2 + 1)\tau \geq l^2\tau$$

for all $i \in \{0, \ldots, l\}$. This implies that $A_1, \ldots, A_l$ can be chosen as stated.
We now choose paths $P_i, \ldots, P_l$ in order, as follows: For $i \in \{1, \ldots, l\}$, we suppose that $P_i, \ldots, P_{i+1}$ have been defined already. To choose $P_i$, we let $C'_i = C_i \setminus N^r(V(P_{i+1}) \cup \cdots \cup V(P_l))$. Since $|V(P_{i+1}) \cup \cdots \cup V(P_l)| \leq l^2 - l$, it follows that $\chi(C'_i) \geq \chi(C_i) - (l^2 - l)\tau \geq l\tau$. Let $C''_i$ be a connected component of $G|\overline{C_i'}$ such that $\chi(C''_i) = \chi(C_i')$. Let $w \in N^r(v) \cap N^{r-1}(A_i)$ such that $w$ has a neighbor in $V(C''_i)$. Then, by Lemma 2.1.1 applied to $G|\{w\} \cup C''_i$ and $w$, there is an induced path $P$ of length at least $\chi(C''_i)/\tau \geq l$ in $G|\{w\} \cup C''_i$ starting at $w$. Let $Q$ be a shortest path from $w$ to $A_i$; then $Q$ contains $r - 1$ vertices. Finally, we let $P_i$ consist of the first $l$ vertices of the concatenated path $QwP$.

For all $i, j \in \{1, \ldots, l\}$ with $i < j$, the sets $V(P_i)$ and $V(P_j)$ are disjoint and $V(P_i)$ is anticomplete to $V(P_j)$.

(2.1)

Let $P$, $Q$ and $w$ be the paths and vertex we chose when constructing $P_i$. By construction, every vertex in $V(P) \setminus \{w\}$ has distance at least $r + 1$ from $P_j$. It follows that $w$ has distance at least $r$ from $P_j$. For all $k \in \{1, \ldots, r\}$, the triangle inequality implies that the vertex in $N^k(v) \cap V(P_i)$ has distance at least $r - (r - k)$ from every vertex in $V(P_j)$. Therefore, the only vertex in $V(P_i)$ that is not guaranteed to have distance at least two from every vertex in $V(P_j)$ is the vertex $z \in V(P_i) \cap N(v)$. Since $z \in N^1(v) \cap V(P_i)$, it follows that $z$ has distance at least $r - (r - 1) = 1$ from every vertex in $V(P_j)$.

Since $z$ is a neighbor of $v$, it follows that the only possible neighbors of $z$ in $V(P_j)$ are in $A_j$ and $N(A_j) \cap N^2(v)$. Since $B$ is stable and $\{z\} \cup A_j \subseteq B$, it follows that $z$ has no neighbor in $A_j$. Suppose that $z$ is adjacent to the vertex $y \in V(P_j) \cap N^2(v)$. By the choice of $P_j$, there exists a path $Q'$ of length $r$ from a vertex in $V(P_j) \cap A_j$ to a vertex $x \in C_j$, and $y \in V(Q') \cap N^2(v)$. Since $z - yQ'x$ is a path of length $r$, it follows that $x$ has distance $r$ from $A_i$. By the choice of $C_j$, it follows that $v$ has
distance \( r + 1 \) from \( x \). But since \( x \in N^r(A_i) \), it follows that \( x \in C_i \), and thus \( x \notin C_j \).

This is a contradiction, and therefore, (2.1) is proved.

Now it follows from (2.1) that \( G|\{(v) \cup V(P_1) \cup \cdots \cup V(P_l)\} \) is isomorphic to an \( l \)-star in \( G \) with center \( v \), as claimed. \( \square \)

2.2 The 2-controlled case

We fix \( l > 0 \) such that \( T \) is an induced subgraph of the graph constructed from two \( l \)-stars by joining their centers by an edge. In this section, we deal with the case that \( G \) is a \( T \)-free \((2, \varphi)\)-controlled graph for some function \( \varphi \). We begin with the following lemma.

**Lemma 2.2.1.** Let \( \kappa \geq 2 \), and let \( G \) be a \( T \)-free graph, and let \( \omega(G) \leq \kappa \). Let \( \tau > 0 \) such that \( \chi(H) \leq \tau \) for every induced subgraph \( H \) of \( G \) with \( \omega(H) \leq \kappa - 1 \).

Then, for every \( v \in V(G) \), there exists \( B \subseteq N(v) \) and \( C \subseteq N(B) \cap N^2(v) \) such that for every vertex \( w \in B \),

\[
\chi(N(N(w) \cap C) \setminus N(v)) \leq (l + 1)^3 \tau^2,
\]

and such that \( \chi(C) \geq \frac{\chi(N^2(v))}{\tau} - (l + 1)^3(\tau^2 + \tau) \).

**Proof.** Let \( G, v, \kappa, \tau \) as in the statement of the lemma. Let \( B' = N(v) \) and \( C' = N^2(v) \). Since \( \omega(G|B') \leq \omega(G) - 1 \leq \kappa - 1 \), it follows that \( \chi(B') \leq \tau \). Thus there exists a stable set \( B \subseteq B' \) such that \( \chi(N(B) \cap C') \geq \chi(C')/\tau \). We write \( C = N(B) \cap C' \).

If \( \chi(C) \leq (l + 1)^3(\tau^2 + \tau) \), then the statement of the lemma holds with \( B = N(v) \) and \( C = \emptyset \). Therefore, we may assume that \( \chi(C) > (l + 1)^3(\tau^2 + \tau) \).

Let \( B_1 \subseteq B \) be maximal with respect to inclusion subject to \( \chi(C \setminus N(B_1)) > (l + 1)^3 \tau^2 \). Let \( C_1 = N(B_1) \cap C \). Since \( \chi(N(w)) \leq \tau \) for all \( w \in V(G) \), it follows that \( \chi(C \setminus C_1) \leq (l + 1)^3 \tau^2 + \tau \). By Lemma 2.1.2 applied to \( G|\{(v) \cup (B \setminus B_1) \cup (C \setminus C_1)\} \), it
follows that there is an induced $l$-star $H$ centered at $v$ in $G|(\{v\} \cup (B \setminus B_1) \cup (C \setminus C_1))$.

By construction, $V(H) \setminus \{v\}$ is anticomplete to $B_1$; we let $C'_1 = C_1 \setminus N(V(H))$. It follows that

$$\chi(C'_1) \geq \chi(C_1) - (l^2 + 1)\tau \geq \chi(C) - (l + 1)^3(\tau^2 + \tau) \geq \frac{\chi(N^2(v))}{\tau} - (l + 1)^3(\tau^2 + \tau).$$

We claim that $B_1$ and $C'_1$ are the desired sets. To prove the claim, we need to prove that for every vertex $w \in B_1$, $\chi(N(N(w) \cap C'_1) \setminus N(v)) \leq (l + 1)^3\tau^2$. Suppose that there is a vertex $w$ for which this does not hold. We apply Lemma 2.1.2 to the vertex $w$ and $G \setminus (V(H) \cup N(V(H) \setminus \{v\}))$. It follows that there is an induced $l$-star $H'$ in $G \setminus (V(H) \cup N(V(H) \setminus \{v\}))$ with center $w$. But since $v$ is adjacent to $w$, it follows that $H' \cup H$ is a double star formed by joining the centers of two $l$-stars, and thus $G$ is not $T$-free, a contradiction. This proves the claim, and concludes the proof of the lemma.

The next lemma will be useful in this section and in Section 2.3; it shows that if $G$ is $(r, \varphi)$-controlled for some $r \in \mathbb{N}$ and function $\varphi$, then there exist two adjacent vertices whose $r$th neighborhoods each have big chromatic number.

**Lemma 2.2.2.** Let $G$ be a graph, and let $\kappa, \tau \in \mathbb{N}$. Suppose that $\omega(G) \leq \kappa$, and for every induced subgraph $H$ of $G$ with $\omega(H) < \kappa$, we have $\chi(H) \leq \tau$. Suppose further that $G$ is $(r, \varphi)$-controlled for some function $\varphi$ and some $r \geq 2$, and that $\chi(N^{r-1}(v)) \leq \tau$ for all $v \in V(G)$.

Let $\tau' > 0$ be such that $\chi(G) > \varphi(\tau') + \tau$. Then there exist $v_1, v_2$ adjacent and $B_1, B_2$ such that $B_i \subseteq N(v_i) \setminus N(v_{3-i})$ for $i = 1, 2$ and such that

$$\chi(N^{r-1}(B_1) \cap N^r(v_1)), \chi(N^{r-1}(B_2) \cap N^r(v_2)) \geq \frac{\tau' - \kappa\tau}{2\kappa}.$$
Proof. We call \( v \in V(G) \) rich if \( \chi(N^r(v)) \geq \tau' \). Let \( Z \) be the set of rich vertices. Since \( G \) is \((r, \varphi)\)-controlled, it follows that \( \chi(G \setminus Z) \leq \varphi(\tau') \), and hence \( \chi(Z) \geq \chi(G) - \varphi(\tau') > \tau \). Consequently, \( \omega(G\vert Z) > \kappa - 1 \). Let \( X \subseteq Z \) be a \( \kappa \)-clique. For each \( Y \subseteq X \), we let \( A(Y) = \{ v \in N(X) : N(v) \cap X = Y \} \).

We say \( Y \) is good if \( \chi(N^{r-1}(A(Y)) \cap N^r(X)) \geq (\tau' - \kappa \tau) / 2^\kappa \). Since every vertex \( x \in X \) is rich, it follows that \( \chi(N^r(x)) \geq \tau' \) for all \( x \in X \). Since

\[
\tau' \leq \chi(N^r(x)) \leq \chi(N^{r-1}(X)) + \sum_{Y \subseteq X, x \in Y} \chi(N^{r-1}(A(Y)) \cap N^r(X)) \\
\leq \kappa \tau + 2^\kappa \max_{Y \subseteq X, x \in Y} \chi(N^{r-1}(A(Y)) \cap N^r(X)),
\]

it follows that for every \( x \in X \), there exists a set \( Y \ni x \) such that \( Y \) is good. Since \( X \) is a maximum clique, it follows that \( A(X) = \emptyset \), and therefore, \( X \) is not a good set. Let \( X_1 \) be a good set with \(|X_1|\) maximum. Since \( X_1 \neq X \), it follows that there exists a vertex \( y \in X \setminus X_1 \). Let \( X_2 \) be a good set containing \( y \). Since \(|X_1| \geq |X_2|\), and since \( y \in X_2 \setminus X_1 \), it follows that \( X_1 \not\subseteq X_2 \) and \( X_2 \not\subseteq X_1 \). Let \( v_1 \in X_1 \setminus X_2, v_2 \in X_2 \setminus X_1 \), and \( B_1 = A(X_1), B_2 = A(X_2) \). Then \( v_1, v_2, B_1, B_2 \) have the desired properties. \( \square \)

We are now ready to prove the main result of Section 2.2.\footnote{2.2}

Lemma 2.2.3. Let \( G \) be a graph, and let \( \kappa, \tau \in \mathbb{N} \). Suppose that \( \omega(G) \leq \kappa \), and for every induced subgraph \( H \) of \( G \) with \( \omega(H) < \kappa \), we have \( \chi(H) \leq \tau \). Then, for every non-decreasing function \( \varphi \), there is a constant \( c \) such that every \((2, \varphi)\)-controlled \( T \)-free graph \( G \) with \( \omega(G) = \kappa \) has \( \chi(G) \leq c \).

Proof. Let \( G, \kappa, \tau \) as in the statement of the lemma, and let \( \varphi : \mathbb{N} \rightarrow \mathbb{N} \) be a non-decreasing function such that \( G \) is \((2, \varphi)\)-controlled. Let

\[
\tau' = 2^\kappa((l + 1)^4 (2 \tau^3 + 2 \tau^2 + \tau) + \tau \varphi(2(l + 1)^3 \tau^2 + \tau) + (\kappa + 1) \tau).
\]

We claim that \( c = \varphi(\tau') + \tau \) works; therefore, we may assume that \( \chi(G) > \varphi(\tau') + \tau \).\footnote{20}
Let \( v_1, v_2, B_1, B_2 \) as in as in Lemma 2.2.2. Let \( k = (\tau' - \kappa \tau)/2^\kappa \); then, by Lemma 2.2.2 it follows that \( \chi(N(B_1)), \chi(N(B_2)) \geq k \). Let \( B'_1 \subseteq B_1 \) be minimal with respect to inclusion such that

\[
\chi(N(B'_1) \setminus (N(v_1) \cup N(v_2))) \geq (l + 1)^4(\tau^2 + \tau);
\]

we let \( C_1 = N(B'_1) \setminus (N(v_1) \cup N(v_2)) \). Let \( C_2 = N(B_2) \setminus (N(B'_1) \cup N(v_1) \cup N(v_2)) \); it follows that

\[
\chi(C_2) > k - (l + 1)^4(\tau^2 + \tau) - \tau \geq (l + 1)^4(2\tau^3 + \tau^2) + \tau \varphi(2(l + 1)^3\tau^2 + \tau).
\]

Moreover, \( B'_1 \) is anticomplete to \( C_2 \).

We now apply Lemma 2.2.1 to \( v_2 \) and \( G' = G | (\{v_1, v_2\} \cup B'_1 \cup B_2 \cup C_1 \cup C_2) \). It follows that there exist \( B'_2 \subseteq B_2 \), and \( C'_2 \subseteq C_2 \), such that \( \chi(N(w) \cap C'_2) \cap (V(G) \setminus B_2) \leq (l + 1)^3\tau^2 \) and \( C'_2 \subseteq N(B'_2) \) and

\[
\chi(C'_2) > \chi(C_2)/\tau - (l + 1)^4(\tau^2 + \tau) \geq (l + 1)^4\tau^2 + \varphi(2(l + 1)^3\tau^2 + \tau).
\]

We call a vertex \( w \) in \( B'_2 \) big if \( \chi(N(w) \cap B'_1) \cap C_1) \geq (l + 1)^3(\tau^2 + \tau) \). Let \( B''_2 \) be the set of big vertices in \( B'_2 \).

We claim the following:

\[
(2.2) \quad \chi(C'_2 \setminus N(B''_2)) \leq (l + 1)^3\tau^2.
\]

Suppose not. We apply Lemma 2.1.2 to \( v_2 \) and \( G | (\{v_2\} \cup (B'_2 \setminus B''_2) \cup C'_2) \); it follows from Lemma 2.1.2 that there is an induced subgraph \( H \) of \( G \) such that \( H \) is an \( l \)-star centered at \( v_2 \) and \( V(H) \subseteq \{v_2\} \cup (B'_2 \setminus B''_2) \cup C'_2 \). Let \( C'_1 = C_1 \setminus \)
\((N(N(V(H)) \cap B'_1) \cup (N(V(H)) \cap C_1))\). Since \(B'_1\) is anticomplete to \(C_2\) and \(v_2\), it follows that every vertex in \(C_1 \setminus C'_1\) either is a neighbor of \(H\) or has a neighbor in \(B'_1\) with a neighbor in \(B'_2\). Since \(V(H) \cap B''_2 = \emptyset\), and since \(|V(H) \cap B'_1| \leq l\), it follows that \(\chi(C_1 \setminus C'_1) \leq l(l + 1)^3(\tau^2 + \tau) + l^2 \tau\). This implies that \(\chi(C'_1) \geq (l + 1)^3\tau^2\), and therefore, by Lemma 2.1.2 applied to \(v_1\) and \(G|\{v_1\} \cup (B'_1 \setminus N(V(H)) \cup C'_1)\), it follows that there exists an induced subgraph \(H'\) of \(G\) such that \(H'\) is an \(l\)-star centered at \(v_1\) and \(V(H)\) and \(V(H')\) are anticomplete except for the edge \(v_1v_2\). This is a contradiction since \(G\) is \(T\)-free and \(G|\{V(H) \cup V(H')\}\) contains \(T\). This proves (2.2).

\[(2.3)\quad \chi(N(B''_2) \cap C'_2) \leq \varphi(2(l + 1)^3\tau^2 + \tau).\]

Suppose not. Let \(C''_2 = N(B''_2) \cap C'_2\) and \(G' = G|C''_2\). Since \(G\) is \((2, \varphi)\)-controlled, it follows that there is a vertex \(w \in C''_2\) such that \(\chi(N_{G'}(w)) > 2(l + 1)^3\tau^2 + \tau\). Let \(B_3 = N(w) \cap C''_2\) and \(C_3 = N_{G'}^2(w)\). Let \(u\) be a neighbor of \(w\) in \(B''_2\). We let \(B'_3 = B_3 \setminus N(u)\) and \(C'_3 = (C_3 \setminus N(u)) \cap N(B'_3)\).

Since \(\chi(N(N(u) \cap C''_2) \cup (V(G) \setminus B'_2)) \leq (l + 1)^3\tau^2\) by our construction and since \(\chi(N(u) \cap C'_2) \leq \tau\), it follows that \(\chi(C'_3) > (l + 1)^3\tau^2\). We apply Lemma 2.1.2 to \(G'|\{w\} \cup B'_3 \cup C'_3\); it follows that there is an induced subgraph \(H\) of \(G\) which is an \(l\)-star centered at \(w\) and \(V(H) \subseteq \{w\} \cup B'_3 \cup C'_3\).

It follows that \(N(u) \cap V(H) = \{w\}\). Let \(B''_1 = B'_1 \cap N(u)\) and let \(C'_1 = C_1 \cap N(B'_1)\). Since \(u \in B''_2\), it follows that \(u\) is big, and so \(\chi(C'_1) \geq (l + 1)^3(\tau^2 + \tau)\). Moreover, \(B''_1\) is anticomplete to \(V(H)\), since \(V(H) \subseteq C'_2\). Let \(C''_1 = C'_1 \setminus (N(V(H)) \cup N(u))\). It follows that \(\chi(C''_1) \geq (l + 1)^3\tau^2\). We apply Lemma 2.1.2 to \(G'|\{u\} \cup B''_1 \cup C''_1\). It follows that there is an induced \(l\)-star \(H'\) centered at \(u\) in \(G'|\{u\} \cup B''_1 \cup C''_1\). But then \(G|\{V(H) \cup V(H')\}\) contains \(T\), a contradiction. This proves (2.3).
Together, (2.2) and (2.3) imply that \( \chi(C'_2) \leq (l + 1)^4 \tau^2 + \varphi(2(l + 1)^3 \tau^2 + \tau) \) contrary to our assumption; this in turn implies that \( \chi(G) \leq \varphi(\tau') + \tau \), which proves the lemma.

Lemma 2.2.3 implies that for every non-decreasing function \( \varphi \), we may assume that there is a \( T \)-free graph that is not \( (2, \varphi) \)-controlled, for otherwise Theorem 2.0.3 holds.

2.3 The \( r \)-controlled case

We now consider the case that there is a function \( \varphi \) such that \( G \) is \( (r, \varphi) \)-controlled for some \( r \geq 3 \). By choosing \( r \) minimal with this property, and by the results of the previous section, we can now assume that \( G \) is not \( (r - 1, \varphi) \)-controlled.

Lemma 2.3.1. Let \( r \geq 3 \) be fixed. Let \( G \) be a \( T \)-free graph, and let \( \tau \in \mathbb{N} \) such that \( \chi(N^{r-1}[v]) \leq \tau \) for all \( v \in V(G) \).

Let \( \tau' \in \mathbb{N} \). Let \( v_1, v_2 \in V(G) \) be adjacent, and for all \( i \in \{1, 2\} \), let \( A_i \subseteq N(v_i) \setminus N(v_{3-i}) \) such that \( \chi(N^r(v_i) \cap N^{r-1}(A_i)) \geq \tau' \). Then there is a vertex \( u \in A_1 \) such that

\[
\chi(N^r(u) \cap N^{r-1}(N(u) \cap A_2) \cap N^r(v_2)) \geq \frac{\tau' - 2(l + 1)^3 \tau^2 - (l + 1)^2 \tau}{l}.
\]

Proof. Suppose not. Let \( A'_2 \subseteq A_2 \) be maximal such that

\[
\chi((N^{r-1}(A_1) \cap N^r(v_1)) \setminus (N^{r-1}(A'_2) \cap N^r(v_2))) > (l + 1)^3 \tau^2.
\]

For \( i \in \{1, \ldots, r\} \), we let \( P_i = N^{i-1}(A'_2) \cap N^i(v_2) \). It follows that \( P_1 = A'_2 \), and that \( \chi(P_r) \geq \tau' - (l + 1)^3 \tau^2 - \tau \).

For \( i \in \{1, \ldots, r\} \), we let \( L_i = N^i(v_1) \cap N^{i-1}(A_1) \). Let \( L'_r = L_r \setminus P_r \). It follows that \( \chi(L'_r) > (l + 1)^3 \tau^2 \). For \( i \in \{1, \ldots, r - 1\} \) in order from \( r - 1 \) down to 1, we
let $L'_i = L_i \cap N(L_{i+1})$. It follows that $L'_r = N^r_{G((\{v_1\} \cup L_1 \cup \cdots \cup L_r)}(v_1)$. By Lemma 2.1.2 applied to $G|((\{v_1\} \cup L_1 \cup \cdots \cup L_r)$, it follows that there is an induced $l$-star $H$ centered at $v_1$ in $G|\{N^r[v_1] \cap N^{r-1}[A_1]\}$.

We let $P'_r = P_r \setminus N^{r-1}(V(H))$. It follows that $\chi(P'_r) \geq r' - (l + 1)^3 r^2 - (l + 1)^2 \tau$. For $i \in \{1, \ldots, r - 1\}$ in order from $r - 1$ down to 1, we let $P'_i = P_i \cap N(P_{i+1})$. Since every vertex in $P'_r$ has distance at least $r$ from $V(H)$, and every vertex in $P'_2 \cup P'_3 \cup \cdots \cup P'_r$ has distance at most $r - 2$ from $P'_r$, the triangle inequality implies that every vertex in $P'_2 \cup P'_3 \cup \cdots \cup P'_r$ has distance at least two from $V(H)$. Therefore, $N(V(H)) \cap (\{v_2\} \cup P'_1 \cup \cdots \cup P'_r) \subseteq P'_1$. By definition, every vertex in $L'_i$ has distance at least $r + 1$ from $v_2$. This implies that every vertex in $L'_2 \cup \cdots \cup L'_r$ has distance at least three from $v_2$, and therefore, $L'_2 \cup \cdots \cup L'_r$ is anticomplete to $P'_1$. It follows that the only possible edges between $V(H)$ and $P'_1$ are between $L'_1$ and $P'_1$.

Since $|V(H) \cap N(v_1)| = l$, it follows that $|V(H) \cap L'_1| \leq l$. Let

$$P''_r = P'_r \setminus \left( \bigcup_{u \in V(H) \cap L'_1} N^{r-1}(N(u) \cap P'_2) \right).$$

Since we may assume that the statement of the lemma does not hold, it follows that for all $u \in V(H) \cap L'_1$,

$$\chi(P'_r \cap N^{r-1}(N(u) \cap P'_2)) \leq \frac{r' - 2(l + 1)^3 r^2 - (l + 1)^2 \tau}{l}.$$

This implies that $\chi(P''_r) \geq (l + 1)^3 r^2$. For $i \in \{1, \ldots, r - 1\}$ in order from $r - 1$ down to 1, we let $P''_i = P'_i \cap N(P''_{i+1})$. Every vertex in $P''_r$ has distance at least $r + 1$ from every vertex $V(H) \cap L'_i$, and therefore, it follows that every vertex in $P''_1$ has distance at most $r - 1$ from $P''_r$. Using the triangle inequality, this implies that every vertex in $P''_1$ has distance at least two from $V(H) \cap L'_1$, and so $V(H)$ is anticomplete to $P''_1 \cup \cdots \cup P''_r$. 

24
By Lemma 2.1.2, it follows that $G|((v_2) \cup P_1' \cup \cdots \cup P''_r)$ contains an induced $l$-star $H'$ centered at $v_2$; and thus $G|(V(H) \cup V(H'))$ contains $T$ as an induced subgraph. This is a contradiction, and the statement of the lemma follows.

Lemma 2.3.2. Let $r \geq 3$ be fixed. Let $G$ be a $T$-free graph, and let $\tau \in \mathbb{N}$ such that $\chi(N^{r-1}[v]) \leq \tau$ for all $v \in V(G)$. Let $\kappa \in \mathbb{N}$ such that $\omega(G) \leq \kappa$, and suppose that $\chi(H) \leq \tau$ for all induced subgraphs $H$ of $G$ with $\omega(H) < \kappa$.

Suppose that $G$ is $(r, \varphi)$-controlled for some function $\varphi$. Then, for every ordering of $V(G)$, some vertex $v$ satisfies

$$\chi(N^{r}(v) \cap N^{r-1}(F(v))) \geq \frac{\varphi^{-1}(\chi(G) - \tau) - \kappa \tau - 1}{3 \cdot 2^{\kappa} - 3(l + 1)^{3} \cdot \tau^{2}},$$

where $F(v)$ denotes those neighbors of $v$ that appear after $v$ in the ordering of $V(G)$.

Proof. Suppose not; and let $\preceq$ be an ordering of $V(G)$, and let $v_1, v_2, B_1, B_2$ be as in Lemma 2.2.2. It follows that

$$\chi(N^{r-1}(B_i) \cap N^{r}(v_i)) \geq (\varphi^{-1}(\chi(G) - \tau) - \kappa \tau - 1)/2^{\kappa}$$

for $i \in \{1, 2\}$. Let $\tau' = (\varphi^{-1}(\chi(G) - \tau) - \kappa \tau - 1)/2^{\kappa}$. Let $w \in B_1 \cup B_2$ be maximal (with respect to the ordering $\preceq$ of $V(G)$) such that for $B = \{u \in B_1 \cup B_2 : u \preceq w\}$,

$$\chi((N^{r}(v_1) \cap N^{r-1}(B \cap B_1)) \cup (N^{r}(v_2) \cap N^{r-1}(B \cap B_2)) \geq 2\tau'/3.$$ 

It follows that there exists an $i \in \{1, 2\}$ such that $\chi((N^{r}(v_i) \cap N^{r-1}(B \cap B_i)) \geq \tau'/3$. By symmetry, we may assume that $i = 1$. By the choice of $B$, it follows that $\chi((N^{r}(v_2) \cap N^{r-1}(B \cap B_2)) \geq \tau'/3 - \tau$. 

25
But now Lemma 2.3.1 applied to $v_1, v_2, B_1 \cap B$ and $B_2 \setminus B$ implies that there is a vertex $u \in B_1 \cap B$ such that

$$
\chi(N^r(u) \cap N^{r-1}(N(u) \cap (B_2 \setminus B))) \geq (\tau'/3 - 3(l+1)^3/3)/l.
$$

But every vertex in $x \in B_2 \setminus B$ satisfies that $u \preceq w \prec x$, and thus $u$ is the desired vertex.

Let $G$ be a graph, and $A, B \subseteq V(G)$ be disjoint sets. We say that $A$ covers $B$ if $B \subseteq N(A)$.

An $r$-leveling $\mathcal{L} = (v, L_1, \ldots, L_r)$ in a graph $G$ consists of a vertex $v \in V(G)$ and for all $i \in \{1, \ldots, r\}$, a set $L_i$ such that the following hold:

- for all $i \in \{1, \ldots, r\}$, we have that $L_i \subseteq N^i(v)$;
- $L_1 \subseteq N(v)$, and for all $i \in \{1, \ldots, r-1\}$, the set $L_i$ covers $L_{i+1}$;
- for all $i, j \in \{1, \ldots, r\}$ with $|i - j| \geq 2$, $L_i$ is anticomplete to $L_j$; and
- $v$ is anticomplete to $L_2 \cup \cdots \cup L_r$.

The sets $L_1, \ldots, L_r$ are called layers of the leveling. An $r$-leveling with an $l$-spire is an $(r+l)$-leveling $\mathcal{L} = (v, \{v_2\}, \ldots, \{v_l\}, \{v\}, L_1, \ldots, L_r)$. The path with vertex set $v_1, \ldots, v_l$ is called the spire of $\mathcal{L}$.

An $r$-leveling with an $l$-spire $\mathcal{L} = (v, \{v_2\}, \ldots, \{v_l\}, \{v\}, L_1, \ldots, L_r)$ is $\tau$-conservative if $\chi(L_r \cap N^r(w) \cap N^{r-1}(N(w) \cap L_r)) \leq \tau$ for all $w \in L_{r-1}$.

The following lemma gives sufficient conditions for the existence of $r$-levelings, $r$-levelings with $l$-spires, and $\tau$-conservative $r$-levelings with $l$-spires.

**Lemma 2.3.3.** Let $\varphi$ be a non-decreasing function, let $r > 2$ and $\tau > 0$. Let $G$ be a $T$-free $(r, \varphi)$-controlled graph such that $\chi(N_r[v]) \leq \tau$ for all $v \in V(G)$. Then $G$ contains an $r$-leveling $\mathcal{L} = (\{v\}, L_1, \ldots, L_r)$ with $\chi(L_r) \geq \varphi^{-1}(\chi(G)) - \tau$. 

26
Suppose further that $\mathcal{L} = \{v\}, L_1, \ldots, L_r$ is an $r$-leveling in $G$. Then $G$ contains an $r$-leveling with an $l$-spire $\mathcal{L}' = (v_1, \{v_2\}, \ldots, \{v_l\}, \{v\}, L'_1, \ldots, L'_r)$ with $L'_i$ stable and $\chi(L'_i) \geq \chi(L_r)/\tau - (2l + 1)\tau$. Moreover, if $\mathcal{L}$ is $\tau'$-conservative for some $\tau' \in \mathbb{N}$, then so is $\mathcal{L}'$.

If $G$ is a $T$-free $(r, \phi)$-controlled graph such that $\chi(N^{r-1}[v]) \leq \tau$ for all $v \in V(G)$, then $G$ contains an $r$-leveling with an $l$-spire $\mathcal{L} = (v_1, \{v_2\}, \ldots, \{v_l\}, \{v\}, L'_1, \ldots, L'_r)$ with $L'_i$ stable and

$$
\chi(L'_i) \geq \phi^{-1}(\chi(G))/\tau - (2l + 1)\tau - 1.
$$

Proof. Let $G, r, \phi, \tau$ be as in the lemma. We begin by proving the first statement of the lemma. Since $G$ is $(r, \phi)$-controlled, it follows that there is a vertex $w \in V(G)$ such that $\chi(G) \leq \phi(\chi(N^r[w]))$. For $i \in \{1, \ldots, r\}$, we let $L_i = N^i(w)$. Then $(\{w\}, L_1, \ldots, L_r)$ is an $r$-leveling. Moreover, since $\chi(N^{r-1}[w]) \leq \tau$, it follows that $\chi(L_r) \geq \chi(N^r[w]) - \tau \geq \phi^{-1}(\chi(G)) - \tau$, as claimed. This proves the first statement of the lemma.

For the second statement, we let $(\{w\}, L_1, \ldots, L_r)$ be as in the lemma. We may assume that $V(G) = \{w\} \cup L_1 \cup \cdots \cup L_r$ by deleting all other vertices. This implies that $N^i(w) = L_i$ for all $i \in \{1, \ldots, r\}$.

Since $L_1 \subseteq N(w)$, it follows that $\chi(L_1) \leq \tau$. Therefore, there is a partition of $L_1$ into $\tau$ stable sets $B_1, \ldots, B_\tau$. Since $L_r \subseteq N^{r-1}(L_1)$, it follows that there is an $i \in \{1, \ldots, \tau\}$ such that $\chi(L_r \cap N^{r-1}(B_i)) \geq \chi(L_r)/\tau$. Choose such an $i$; and let $B = B_i$ and $C = L_r \cap N^{r-1}(B_i)$.

Next, we let $A \subseteq B$ be maximal with respect to inclusion subject to the condition that $\chi(C \setminus N^{r-1}(A)) \geq l\tau$. It follows that $\chi(N^{r-1}(A) \cap C) \geq \chi(C) - (l + 1)\tau$. Now let $v \in N(C \setminus N^{r-1}(A)) \cap L_{r-1}$ such that $v$ is in a connected component $K$ of $G|(N(C \setminus N^{r-1}(A)) \cap L_{r-1})$ with chromatic number at least $l\tau$. Let $P$ be a path of length $r - 2$ from $v$ to a vertex $a \in B \setminus A$. 

27
By Lemma 2.1.1 and since $\chi(C \setminus N^{r-1}(A)) \geq l\tau$, it follows that $G[(\{v\} \cup K)$ contains an induced path $Q$ of length at least $l$ starting at $v$. Let $v' \in N(v) \cap V(Q)$; it follows that $v' \in C \setminus N^{r-1}(A)$. Let $P'$ be the path consisting of the first $l$ vertices of $aPvQ$; we write $a = v_l, v_{l-1}, \ldots, v_1$ for the vertices of $P'$ in order.

Now let $H = G \setminus (V(P') \cup N(V(P')))$. We let $L'_1 = A$ and for $i \in \{2, \ldots, r\}$, we let $L'_i = N^i_H^{-1}(A) \cap N^i_H(w)$. It follows that $L'_i \subseteq L_i$ for all $i \in \{1, \ldots, r\}$. Therefore,

$$L' = (v_1, \{v_2\}, \ldots, \{v_r\}, L'_1, \ldots, L'_r)$$

is an $r$-leveling with an $l$-spire in which the set $L'_1$ is stable. Moreover, since $L'_{r-1} \subseteq L_{r-1}$ and $L'_r \subseteq L_r$, it follows that if $L$ is $\tau'$-conservative for some $\tau' \in \mathbb{N}$, then $L'$ is $\tau'$-conservative as well.

We claim that $(C \cap N^{r-1}_G(A)) \setminus N^{r-1}_G[V(P')] \subseteq L'_r$. Suppose not; and let $c \in (C \cap N^{r-1}_G(A)) \setminus N^{r-1}_G[V(P')]$ such that $c \notin L'_r$. Let $Q'$ be a path of length $r - 1$ with ends $c$ and $a'$ for some $a' \in A$. Then, since $c \notin L'_r$, it follows that $V(Q') \notin V(H)$.

Since $c \notin N^{r-1}_G[V(P')]$, it follows that $c$ has distance at least $r$ from $P'$, and so $c \in V(H)$. Moreover, for all $i \in \{1, \ldots, r\}$, the vertex in $N^{i-1}(A) \cap N^i(w)$ has distance at most $r - i$ from $c$, and by the triangle inequality, it has distance at least $i$ from $V(P')$. It follows that $V(Q') \setminus A \subseteq V(H)$, and that $a'$ has distance one from $V(P')$. Since $B$ is stable, it follows that $a'$ and $a$ are non-adjacent, and thus $a'$ is adjacent to the vertex $b \in N(a) \cap V(P')$. Now $a' - bPv - v'$ is a path from $a' \in A$ to $v'$ of length $r - 1$, contrary to the fact that $v' \in C \setminus N^{r-1}(A)$. This is a contradiction, and thus our claim that $(C \cap N^{r-1}_G(A)) \setminus N^{r-1}_G[V(P')] \subseteq L'_r$ is proved. It follows that $\chi(L'_r) \geq \chi(L_r)/\tau - (2l + 1)\tau$, which completes the proof of the second statement of the lemma.

The third statement of the lemma is follows from the first two statements. This completes the proof. □
The next lemma is used to show that \( \tau \)-conservative levelings are useful for bounding the chromatic number of their layers. Later on, we will prove that in a graph \( G \), under the right assumptions, we can always find a \( \tau \)-conservative leveling whose layers have chromatic number close to \( \chi(G) \), thus providing an upper bound on \( \chi(G) \).

**Lemma 2.3.4.** Let \( \varphi, \psi \) be non-decreasing functions. Let \( r > 2 \) and \( \kappa, \tau \in \mathbb{N} \). Then there exists a constant \( c \) such that the following holds. Let \( G \) be a \( T \)-free \((r, \varphi)\)-controlled graph such that \( \omega(G) \leq \kappa \), and such that \( \chi(N^{r-1}[v]) \leq \tau \) for all \( v \in V(G) \), and \( \chi(H) \leq \tau \) for all induced subgraphs \( H \) of \( G \) with \( \omega(H) < \kappa \).

Suppose that in every induced subgraph \( H \) of \( G \), there exists an \( r \)-leveling with an \( l \)-spire \( L = (v_1, \{v_2\}, \ldots, \{v_l\}, L_1, \ldots, L_r) \) in \( G \) such that \( L \) is \( \tau \)-conservative and \( \psi(\chi(L_r)) \geq \chi(H) \). Then \( \chi(G) \leq c \).

**Proof.** Let \( k' = R(\kappa+1, l) \) be the Ramsey number of \( \kappa+1 \) and \( l \) as in Theorem 1.3.1, and let \( k = R(2^{\kappa^2}k', l) \) be the Ramsey number of \( k' \) and \( l \) as in Theorem 1.3.1. We define a series \( \mathcal{L}_1, \ldots, \mathcal{L}_k \) of \( r \)-levelings with \( l \)-spires as follows:

- \( \mathcal{L}_1 = (v_1, \{v_2\}, \ldots, \{v_l\}, \{v\}, L_1, \ldots, L_r) \) is a \( \tau \)-conservative \( r \)-leveling with an \( l \)-spire in \( G \); and

- For \( i > 1 \), \( \mathcal{L}_i = (v_1, \{v_2\}, \ldots, \{v_l\}, \{v\}, L_1, \ldots, L_r) \) is a \( \tau \)-conservative \( r \)-leveling with an \( l \)-spire in \( G|L^{-1}_i \).

We let \( C = L_r^k \setminus N^2 \left[ \bigcup_{i \in \{1, \ldots, k\}} \{v_i^1, \ldots, v_i^l\} \right] \). Thus, for all \( i \in \{1, \ldots, k\} \), no vertex in \( C \) is within distance two of the spire of \( \mathcal{L}_i \). Moreover, by the definition of the \( \mathcal{L}_i \), it follows that for all \( i, j \in \{1, \ldots, k\} \) with \( i \neq j \), the set \( \{v_i^1, \ldots, v_i^l, v_j^1, v_j^l\} \cup L_1^i \cup \cdots \cup L_r^i \) is anticomplete to \( \{v_i^1, \ldots, v_i^l, v_j^1, v_j^l\} \cup L_1^i \cup \cdots \cup L_r^i \cup C \).

We claim the following.

\[
\chi(C) \leq 2^{k^2}(\tau + \varphi(\kappa \tau + 2^\kappa((l+1)^3 \tau^2 + (2l+1)\tau))).
\]

29
We first prove the statement of the lemma assuming that (2.4) holds. Let $L_r^0 = V(G)$. It follows that for all $i \in \{0, \ldots, k-1\}$, we have that $\chi(L_r^i) \leq \psi(\chi(L_r^{i+1}))$. This implies that

$$\chi(G) \leq \psi^k(L_r^k) \leq \psi^k(\chi(C) + lk\tau) \leq \psi^R(2^{\kappa^2 R(\kappa+1,l)})(2^{k^2}(\tau + \varphi(\kappa\tau + 2^\kappa((l+1)^3\tau^2 + (2l+1+k)\tau)))$$

where $\psi^k$ denotes the $k$-fold application of $\psi$. This implies the statement of the lemma.

We now prove (2.4). Suppose that (2.4) does not hold. For $c \in C$ and $i \in \{1, \ldots, k\}$, it follows that $c$ has a neighbor in $L_r^{i-1}$. For each $c$, we define $p(c,i)$ to be such a neighbor. Let $I = \{(i,j) : i,j \in \{1, \ldots, k\}, i < j\}$. For every $c \in C$, we define a function $f_c : I \rightarrow \{0,1\}$ by setting $f_c((i,j)) = 1$ if and only if $p(c,i)$ has a neighbor in $\{v_i\} \cup L_i^1 \cup \cdots \cup L_i^{r-1}$ (note that $p(c,i)$ is non-adjacent to the spire of $L_i$ by the definition of $C$). Since $|I| \leq k^2$, it follows that there is a set $C'' \subseteq C$ such that $\chi(C'') \geq \chi(C)/2^{k^2}$ and $f_c((i,j)) = f_d((i,j))$ for all $c,d \in C''$ and $(i,j) \in I$. We define $f((i,j)) = f_c((i,j))$ for some $c \in C''$.

Now let $H$ be the graph with vertex set $\{1, \ldots, k\}$ and in which $i,j$ with $i < j$ are adjacent if and only if $f((i,j)) = 1$. By Theorem 1.3.1 it follows that there is a subset $A \subseteq \{1, \ldots, k\}$ such that either

(i) $|A| \geq l$ and $f((i,j)) = 1$ for all $i,j \in A$ with $i < j$; or

(ii) $|A| \geq 2^{\kappa^2}k'$ and $f((i,j)) = 0$ for all $i,j \in A$ with $i < j$. 

30
We consider these cases separately. Suppose first that (i) happens. Let $A = \{i_1, \ldots, i_l\}$ with $i_1 < \cdots < i_l$. We know that $\chi(C') \geq \varphi((l+1)^3 \tau^2 + 2\tau)$. This implies that since $G$ is $(r, \varphi)$-controlled, there is a vertex $v \in C'$ such that $\chi(N^r_{G|C'}(v)) \geq (l+1)^3 \tau^2 + 2\tau$.

Let $w = p(v, i_1)$. By the definition of $A$, it follows that $w$ has a neighbor in $L_1 \cup \cdots \cup L_{r-1}^{i_j}$ for all $j \in \{1, \ldots, l\}$. For $j \in \{1, \ldots, l\}$, we let $P_j$ be an induced path of length $l + 1$ in $\{v_1^{i_j}, \ldots, v_l^{i_j}, v^{i_j}\} \cup L_1^{i_j} \cup \cdots \cup L_{r-2}^{i_j} \cup \{w\}$ with an end in $w$. Such a path exists; we can choose $P_j$ to consists of the first $l + 2$ vertices of a shortest path from $w$ to $v_1^{i_j}$. By construction, it follows that for all $j, j' \in \{1, \ldots, l\}$ with $j < j'$, we have that $V(P_j) \setminus \{w\} \subseteq \{v_1^{i_j}, \ldots, v_l^{i_j}, v^{i_j}\} \cup L_1^{i_j} \cup \cdots \cup L_{r-2}^{i_j}$, and $V(P_{j'}) \setminus \{w\} \subseteq L_{r-2}^{i_j}$, and therefore, $V(P_j) \setminus \{w\}$ is anticomplete to $V(P_{j'}) \setminus \{w\}$. It follows that $H = G|\{\{w\} \cup V(P_1) \cup \cdots \cup V(P_l)\}$ is an induced $l$-star in $G$.

By construction, $V(H) \setminus \{w\}$ is anticomplete to $C'$. Since $G_{i_1}$ is $\tau$-conservative, it follows that $\chi(C' \cap N^r(w) \cap N^{r-1}(N(w) \cap C')) \leq \tau$. Let $G' = G|(C' \setminus (N(w) \setminus \{v\}))$. We claim that $\chi(N^r_{G'}(v)) \geq (l+1)^3 \tau^2$. Suppose not. Since

$$N^r[w] \cap N^{r-1}[N(w) \cap C'] \subseteq N^{r-1}[w] \cup (C' \cap N^r(w) \cap N^{r-1}(N(w) \cap C')),$$

it follows that $\chi(N^r[w] \cap N^{r-1}[N(w) \cap C']) \leq 2\tau$. This implies that there is a vertex $c \in N^r_{G|C'}(v) \setminus (N^r[w] \cap N^{r-1}[N(w) \cap C'])$ such that $c \not\in N^r_{G'}(v)$. Let $P$ be a path from $c$ to $v$ of length $r$ with $V(P) \subseteq C'$. Then $V(P) \not\subseteq V(G')$, and hence $V(P) \setminus \{v\}$ contains a neighbor $a$ of $w$. But then $w-aPc$ is a path of length at most $r-1$ from $w$ to $c$ with $a \in C'$, contrary to the fact that $c \not\in (N^r[w] \cap N^{r-1}[N(w) \cap C'])$. This contradiction proves our claim that $\chi(N^r_{G'}(v)) \geq (l+1)^3 \tau^2$. By Lemma 2.1.2 it follows that there is an $l$-star $H'$ in $G'$ with center $v$. Now $G|(V(H) \cup V(H'))$ contains $T$, a contradiction. This proves that case (i) does not occur.
It follows that case (ii) happens. Let \( \tau' = \kappa \tau + 2^\kappa ((l + 1)^3 \tau^2 + (2l + 1) \tau) \). We call a vertex \( c \in C' \) rich if \( \chi(N_{G|C'}^{r}(v)) \geq \tau' \). Let \( Z \) be the set of rich vertices. Since \( G \) is \((r, \varphi)\)-controlled, it follows that \( \chi(C' \setminus Z) \leq \varphi(\tau') \), and so \( \chi(Z) \geq \chi(C') - \varphi(\tau') > \tau \).

Since \( \chi(H) \leq \tau \) for all induced subgraphs \( H \) of \( G \) with \( \omega(H) < \kappa \), it follows that \( \omega(G|Z) = \kappa \). Let \( K \subseteq Z \) be a \( \kappa \)-clique; we let \( K = \{ x_1, \ldots, x_\kappa \} \).

We define a function \( F : A \rightarrow (2^K)^{\kappa} \) as follows. For all \( j \in A \), we let \( F(j) = (N(p(x_1, j)) \cap K, \ldots, N(x_\kappa, j) \cap K) \). Since \( F(j) \) is a \( \kappa \)-tuple of subsets of \( K \), it follows that there are at most \( 2^{\kappa^2} \) possible values of \( F(j) \). Consequently, there is a set \( A' \subseteq A \) with \( |A'| \geq |A|/2^{\kappa^2} \) and a \( \kappa \)-tuple \( F \) such that \( F(j) = F \) for all \( j \in A' \). Let \( A' = \{ i_1, \ldots, i_\kappa \} \) where \( i_1 < \cdots < i_\kappa \). Let \( F = (Y_1, \ldots, Y_\kappa) \).

For \( Y \subseteq K \), we let \( A(Y) = \{ v \in N(K) : N(v) \cap K = Y \} \). We say \( Y \) is good if \( \chi(N_{G|C'}^{r-1}(A(Y)) \cap N_{G|C'}^{r}(K)) \geq (\tau' - \kappa \tau)/2^\kappa \). Since every vertex \( x \in X \) is rich, it follows that \( \chi(N_{G|C'}^{r}(x)) \geq \tau' \) for all \( x \in X \). Since

\[
\tau' \leq \chi(N_{G|C'}^{r}(x)) \leq \chi(N_{G|C'}^{r-1}(K)) + \sum_{Y \subseteq K \atop x \in Y} \chi(N_{G|C'}^{r-1}(A(Y)) \cap N_{G|C'}^{r}(K)) \\
\leq \kappa \tau + 2^\kappa \max_{Y \subseteq K \atop x \in Y} \chi(N_{G|C'}^{r-1}(A(Y)) \cap N_{G|C'}^{r}(K)),
\]

it follows that for every \( x \in K \), there is a good set \( Y \subseteq K \) with \( x \in Y \). Let \( \mathcal{Y} = \{ Y_1, \ldots, Y_\kappa \} \cup \{ Y \subseteq K : Y \text{ is good} \} \). Let \( Y \in \mathcal{Y} \) be such that \( |Y| = \max \{|Y'| : Y' \in \mathcal{Y}\} \). Since \( K \) is a maximum clique, it follows that \( |Y'| < \kappa \).

We consider two cases. Suppose first that \( Y \in \{ Y_1, \ldots, Y_\kappa \} \). Then we let \( X = Y \). Let \( x \in K \setminus Y \), and let \( X' \subseteq K \) be a good set containing \( x \). Then \( x \in X' \setminus X \), and since \( |X| \geq |X'| \), it follows that \( X \setminus X' \neq \emptyset \). Now consider the case that \( Y \) is a good set. We let \( X' = Y \). Let \( x = x_j \in K \setminus Y \), and let \( X = Y_j \). Then \( x \in X \setminus X' \) and since \( |X'| \geq |X| \), it follows that \( X' \setminus X \neq \emptyset \).

In both cases, we have constructed two sets \( X \) and \( X' \) such that \( X' \) is good, and there exists an \( i \in \{ 1, \ldots, \kappa \} \) such that \( X = Y_i \in \{ Y_1, \ldots, Y_\kappa \} \), and \( X \setminus X', X' \setminus X \neq \emptyset \).
Let \( x \in X \setminus X' \) and \( x' \in X' \setminus X \). For \( j \in A' \), we let \( y_j = p(x, j) \). It follows that \( N(y_j) \cap K = Y_i = X \). Let \( S = \{ y_j : j \in A' \} \). Since \( |S| \geq k' \), it follows from Theorem 1.3.1 that \( G|S \) contains a clique of size \( \kappa + 1 \) or a stable set of size \( l \). Since \( \omega(G) = \kappa \), the former is impossible. It follows that there exists \( S' \subseteq S \) with \( |S'| = l \) and \( S' \) stable. Let \( A'' = \{ j : y_j \in S' \} \). For \( j \in A'' \), we let \( P_j \) be an induced path of length \( l + 1 \) in \( \{ v_1^j, \ldots, v_l^j \} \cup \bigcup_{i=1}^{l} L_i^j \cup \cdots \cup L_{l-2}^j \cup \{ y_i \} \) starting in \( y_i \). Such a path exists, e.g. we can choose the first \( l + 2 \) vertices of a shortest path from \( y_j \) to \( v_1^j \). Then \( G|((\{ x \} \cup (\bigcup_{j \in A''} V(P_j))) \) is an \( l \)-star \( H \). By construction, \( V(H) \setminus \{ x \} \) is anticomplete to \( x' \), and \( V(H) \setminus (\{ x \} \cup \{ y_j : j \in A'' \}) \) is anticomplete to \( C' \).

Let \( B = A(X') \). Since \( X' \) is good, it follows that \( \chi(N_{G|x'}^{-1}(B) \cap N_{r|C'}(x')) \geq (\tau' - \kappa \tau)/2^s \). Moreover, by the definition of \( x, x' \), it follows that \( x' \) is complete to \( B \) and \( x \) is anticomplete to \( B \). Let \( G' = G|(C' \setminus (N(\{ x \} \cup \{ y_j : j \in A'' \}) \setminus \{ x' \})) \).

We claim that \( \chi(N_{C'}(x') \cap N_{r|G'}^{-1}(B)) \geq (l + 1)^3 \tau^2 \). Suppose not. Let \( c \in (N_{G|x'}^{-1}(x') \cap N_{r|G'}^{-1}(B)) \cap (N_{C'}(x') \cap N_{r|G'}^{-1}(B)) \), and let \( P \) be a path of length \( r \) from \( c \) to \( x' \) with \( V(P) \subseteq C' \) such that \( \{ a \} = N(x') \cap V(P) \subseteq B \). It follows that \( V(P) \not\subseteq V(G') \), and hence \( V(P) \setminus \{ x' \} \) contains a neighbor of \( x \) or there is a \( j \in A'' \) such that \( V(P) \setminus \{ x' \} \) contains a neighbor of \( y_j \).

Suppose first that \( V(P) \setminus \{ x', a \} \) contains a neighbor \( b \) of a vertex \( w \) in \( \{ x \} \cup \{ y_j : j \in A'' \} \). Then \( w-b-P-c \) is a path of length at most \( r - 1 \), and so \( c \in N_{r|G'}^{-1}(\{ x \} \cup \{ y_j : j \in A'' \}) \). Let \( C^* = N_{C'}(x') \cap N_{r|G'}^{-1}(\{ x \} \cup \{ y_j : j \in A'' \}) \). It follows that \( \chi(C^*) \leq (l + 1) \tau \), and hence \( \chi(N_{G'}(x') \setminus C^*) \geq \chi(N_{G'}^{-1}(C')) - (l + 1) \tau \).

Now we may assume that \( c \not\in N_{r|G'}^{-1}(\{ x \} \cup \{ y_j : j \in A'' \}) \), and that \( a \) is the only vertex in \( V(P) \setminus V(G') \). Since \( x \) is anticomplete to \( B \) and \( a \in B \), it follows that there exists a \( j \in A'' \) such that \( y_j \) is adjacent to \( a \). Thus \( c \in (C' \cap N_{r}(y_j) \cap N_{r|C'}^{-1}(N(y_j) \cap C')) \). But \( C_{j} \) is \( \tau \)-conservative, and so \( \chi\left( \bigcup_{j \in A'} (C' \cap N_{r}(y_j) \cap N_{r|C'}^{-1}(N(y_j) \cap C')) \right) \leq l \tau \).
It follows that
\[
\chi(N_{G'}(x') \cap N_{G'}^{-1}(B)) \geq \chi((N_{G|C'}(x') \cap N_{G|C'}^{-1}(B))) - (2l + 1)\tau \geq (l + 1)^{3}\tau^2
\]
as claimed. Now Lemma 2.1.2 implies that there is an induced \(l\)-star \(H'\) in \(G'\) with center \(x'\). It follows that \(G|(V(H) \cup V(H'))\) contains \(T\) as an induced subgraph, a contradiction. This proves (2.4), and hence the proof is complete.

A grading of a graph \(G\) is a partition \((W_1, \ldots, W_s)\) of \(V(G)\). It is \(\tau\)-bounded if \(\chi(W_i) \leq \tau\) for all \(i \in \{1, \ldots, s\}\). A 1-bounded grading is called stable. Let \(A\) cover \(B\), and let \(A = \{a_1, \ldots, a_s\}\). Then the \((a_1, \ldots, a_s)\)-grading of \(B\) is the grading \((W_1, \ldots, W_s)\) defined by
\[
\begin{align*}
W_1 &= N(a_1) \cap B \\
W_2 &= (N(a_2) \cap B) \setminus W_1 \\
\vdots \\
W_i &= (N(a_i) \cap B) \setminus (W_1 \cup \cdots \cup W_{i-1}) \\
\vdots \\
W_s &= (N(a_s) \cap B) \setminus (W_1 \cup \cdots \cup W_{s-1}).
\end{align*}
\]
For a grading \(\mathcal{W} = (W_1, \ldots, W_s)\), we let \(\chi(\mathcal{W}) = \chi(W_1 \cup \cdots \cup W_s)\). A grading \((W'_1, \ldots, W'_s)\) is a subgrading of \((W_1, \ldots, W_s)\) if \(W'_i \subseteq W_i\) for all \(i \in \{1, \ldots, s\}\). A refinement of \(\mathcal{W}\) is an ordering of \(W_1 \cup \cdots \cup W_s\) such that if \(i < j\), then every vertex of \(W_i\) appears before every vertex of \(W_j\) in the ordering for all \(i, j \in \{1, \ldots, s\}\).

The next lemma shows how to turn a grading into a stable grading.

**Lemma 2.3.5.** Let \(G\) be a graph with \(\omega(G) \leq \kappa\) and \(\tau > 0\) such that \(\chi(H) \leq \tau\) for all induced subgraphs \(H\) of \(G\) such that \(\omega(H) \leq \kappa - 1\). Let \(A, B \subseteq V(G)\) such that
A = \{a_1, \ldots, a_s\} covers B, and let \( \mathcal{W} \) be the \((a_1, \ldots, a_s)\)-grading of B. Then there is a stable subgrading \( \mathcal{W}' \) of \( \mathcal{W} \) such that \( \chi(\mathcal{W}') \geq \chi(B)/\tau \).

**Proof.** Let \((W_1, \ldots, W_s)\) be the \((a_1, \ldots, a_s)\)-grading of B. For all \( i \in \{1, \ldots, s\} \), we have \( W_i \subseteq N(a_i) \) and so \( \chi(W_i) \leq \tau \) for all \( i \in \{1, \ldots, s\} \). Therefore, for all \( i \in \{1, \ldots, s\} \), \( W_i \) can be partitioned into \( \tau \) stable sets \( W^i_1, \ldots, W^i_\tau \). It follows that there exists a \( j \in \{1, \ldots, \tau\} \) such that \( \chi(W^j_1 \cup \cdots \cup W^j_\tau) \geq \chi(B)/\tau \). It follows that \( \mathcal{W} = (W^j_1, \ldots, W^j_\tau) \) is a stable subgrading of \((W_1, \ldots, W_s)\) with \( \chi(\mathcal{W}) \geq \chi(B)/\tau \) as claimed.

We are now ready to prove the final lemma of this section.

**Lemma 2.3.6.** Let \( \varphi \) be a non-decreasing function. Let \( r > 2 \) and \( \gamma, \tau, \kappa \in \mathbb{N} \). Then there exists a constant \( c \) such that the following holds.

Let \( G \) be a \( T \)-free \((r, \varphi)\)-controlled graph such that

- \( \chi(N^{r-1}[v]) \leq \tau \) for all \( v \in V(G) \);
- \( \omega(G) \leq \kappa \), and \( \chi(H) \leq \tau \) for all induced subgraphs \( H \) of \( G \) with \( \omega(H) < \kappa \); and
- \( G \) does not contain an \( r \)-leveling with an \( l \)-spire

\[
\mathcal{L} = (v_1, \{v_2\}, \ldots, \{v_l\}, \{v\}, L_1, \ldots, L_r)
\]

such that \( \mathcal{L} \) is \((3(l+1)^3\tau^2)\)-conservative and \( \chi(L_r) \geq \gamma \).

Then \( \chi(G) \leq c \).

**Proof.** Let \( G, \varphi, r, \gamma, \tau, \kappa \) be as in the lemma. Let \( \tau' = 12(l+1)^3\tau^2 \). We call a vertex \( v \in V(G) \) rich if \( \chi(N^r(v)) \geq \tau' \). Let \( G' \) be the induced subgraph of rich vertices in \( G \). Since \( G \) is \((r, \varphi)\)-controlled, it follows that \( \chi(G \setminus G') \leq \varphi(\tau') \), and thus \( \chi(G') \geq \chi(G) - \varphi(\tau') \).
By Lemma 2.2.2 it follows that there exist \( v_1, v_2, B_1, B_2 \) such that \( \chi(N_G^r(B_i) \cap N_{G'}^r(v_i)) \geq (\varphi^{-1}(\chi(G') - \tau) - 1 - \kappa \tau)/2^k \) for all \( i \in \{1, 2\} \). Let \( B_1' \subseteq B_1 \) be minimal such that

\[
\chi(N_G^r(B_1') \cap N_{G'}^r(v_1)) \geq (\varphi^{-1}(\chi(G') - \tau) - 1 - \kappa \tau)/2^{k+1} =: \tau^*.
\]

We may assume that \( \tau^* \geq 6l(l + 1)^3 \tau^2 \), for otherwise, the statement of the lemma holds. Let \( B_2' \subseteq B_2 \) be the set of vertices \( u \) in \( B_2 \) such that \( \chi(N^r(u) \cap N^{r-1}(N(u) \cap B_1')) \geq 3(l + 1)^3 \tau^2 \). By Lemma 2.3.1 it follows that \(  \chi(N^r(v_1) \cap N^{r-1}(B_2 \setminus B_2')) < 6l(l + 1)^3 \tau^2 \leq \tau^* \), and thus \( \chi(N^r(v_1) \cap N^{r-1}(B_2')) \geq 2\tau^* - 6l(l + 1)^3 \tau^2 \).

We let

\[
L^1 = (\{v_1\}, L_1, \ldots, L_r)
\]

\[
= (\{v_1\}, B_1, N_G^r(B_1') \cap N_{G'}^r(v_1), \ldots, N_G^r(B_1') \cap N_{G'}^r(v_1))
\]

and

\[
L^2 = (\{v_2\}, L_1, \ldots, L_r)
\]

\[
= (\{v_2\}, B_2, N_G^r(B_2') \cap (N_{G'}^r(v_2) \setminus L_1), \ldots, N_G^r(B_2') \cap (N_{G'}^r(v_2) \setminus L_1)).
\]

It follows that \( L_i^1 \) is anticomplete to \( L_{i+1}^2 \) for all \( i \in \{1, \ldots, r - 1\} \). For \( i \in \{1, \ldots, r\} \), we let \( P_i = N_G^r(\{v_1, v_2\}) \). Since \( \chi(N^{r-1}(w)) \leq \tau \) for all \( w \in V(G) \), it follows that

\[
\chi(L_i^1 \cap P_r) \geq \chi(L_i^1) - \tau \quad \text{for all } i \in \{1, 2\}.
\]

For all \( i \in \{1, 2\} \) and \( j \in \{1, \ldots, r\} \), we let \( Q_j^i = P_j \cap L_j^i \). It follows that for all \( i \in \{1, 2\} \), \( (\{v_1\}, Q_1^i, \ldots, Q_r^i) \) is an \( r \)-leveling. Since \( \chi(N^{r-1}(w)) \leq \tau \) for all \( w \in V(G) \), it follows that \( \chi(L_i^1 \setminus Q_r^i) \leq \tau \) for all \( i \in \{1, 2\} \).
Let $\mathcal{A} = (a_1, \ldots, a_s)$ be an ordering of the vertices in $Q^2_{r-1}$, and let $\mathcal{W}$ be the $\mathcal{A}$-grading of $Q^2_r$. By Lemma 2.3.5, it follows that there is a stable subgrading $\mathcal{W}'$ of $\mathcal{W}$ with $\chi(\mathcal{W}') \geq \chi(Q^2_r)/\tau$. Let $\mathcal{W}' = (W_1, \ldots, W_s)$ and let $B = W_1 \cup \cdots \cup W_s$.

We call a vertex $w \in Q^2_{r-1}$ big if $\chi(N^{-1}(N(w) \cap B) \cap B) \geq 3(l + 1)^3\tau^2$. Let $S$ be the set of vertices in $Q^2_{r-1}$ that are not big. By the assumptions of the lemma, and since $(\{v_2\}, Q^2_1, \ldots, S, Q^2_r \cap N(S))$ is a $(3(l + 1)^3\tau^2)$-conservative $r$-leveling, it follows that $\chi(N(S) \cap Q^2_r) \leq \gamma$. Let $H = G \setminus (N(S) \cap Q^2_r)$. It follows that $\chi(H) \geq \chi(B) - \gamma$.

By Lemma 2.3.2 applied to $H$ and an ordering $\preceq$ that is a refinement of $\mathcal{W}'$, it follows that there is a vertex $w \in V(H)$ such that

$$\chi(N_H^r(w) \cap N_H^{-1}(F(w)) \geq ((\varphi^{-1}(\chi(\mathcal{W}')) - \tau) - \kappa \tau - 1)/(3 \cdot 2^\kappa) - 3(l + 1)^3\tau^2)/l,$$

where $F(v)$ denotes neighbors of $v$ that appear after $v$ with respect to the ordering $\preceq$. Let $i \in \{1, \ldots, s\}$ be such that $w \in W_i$. We may assume that $\chi(N_H^r(w) \cap N_H^{-1}(F(w)) \geq 6(l + 1)^3\tau^2$, for otherwise the statement of the lemma holds.

Since $W_i$ is stable, it follows that $F(w) \subseteq W_{i+1} \cup \cdots \cup W_s$. Consequently, $a_i$ is anticomplete to $F(v)$. Since $a_i \notin S$, it follows that $a_i$ is big, and thus $\chi(N^{-1}(N(a_i) \cap B) \cap B) \geq 3(l + 1)^3\tau^2$.

By Lemma 2.3.1 applied to $a_i$, $w$, $N(a_i) \cap Q^2_{r-2}$, $F(w)$, since $\chi(N_H^r(w) \cap N_H^{-1}(F(w)) \geq 6(l + 1)^3\tau^2$, and since $N(a_i) \cap Q^2_{r-2}$ is anticomplete to $F(w)$, it follows that $\chi(N^{-1}(N(a_i) \cap P_{r-2}) \cap N^r(a_i)) \leq 3(l + 1)^3\tau^2$. We define $w_r = w$ and $w_{r-1} = a_i$. Let $R$ be a path from $w_r$ to $w_1 \in B'_1$ with vertices $w_r, w_{r-1}, \ldots, w_1$ in order and such that $w_j \in Q^2_j$ for all $j \in \{1, \ldots, r\}$. By the definition of $Q^2_j$, it follows that the vertices $w_{r-1}, \ldots, w_1$ are rich.

We claim that for every $j \in \{1, \ldots, r - 2\}$, the following hold:

- $\chi(N^r(w_j) \cap N^{r-1}(N(w_j) \cap P_{j+1})) \geq 3(l + 1)^3\tau^2$;
\[ \chi(N^r(w_j) \cap N^{r-1}(N(w_j) \cap P_{j-1})) < 3(l + 1)^3 \tau^2; \text{ and} \]
\[ \chi(N^r(w_j) \cap N^{r-1}(N(w_j) \cap P_j)) < 6(l + 1)^3 \tau^2. \]

We prove this by induction. Since the first and second bullet hold for \( j = r - 1 \), we may assume that the first and second bullet hold for \( w_{j+1} \). Since \( w_j \) is rich, and since \( N(w_j) \subseteq P_{j+1} \cup P_{j-1} \cup P_j \), it follows that

\[
\chi(N^r(w_j) \cap N^{r-1}(N(w_j) \cap P_{j+1})) \\
+ \chi(N^r(w_j) \cap N^{r-1}(N(w_j) \cap P_{j-1})) \\
+ \chi(N^r(w_j) \cap N^{r-1}(N(w_j) \cap P_j)) \geq 12(l + 1)^3 \tau^2.
\]

Therefore it suffices to prove that the second and third bullet hold for \( w_j \); then the statement of the first bullet follows.

Suppose first that the second bullet does not hold. Since \( \chi(N^r(w_{j+1}) \cap N^{r-1}(N(w_{j+1}) \cap P_{j+2})) \geq 3(l + 1)^3 \tau^2 \), since the first bullet holds for \( w_{j+1} \), and since \( P_{j+2} \) is anticomplete to \( P_{j-1} \), it follows that \( w_j, w_{j+1} \) and the sets \( N(w_j) \cap P_{j-1} \) and \( N(w_{j+1}) \cap P_{j+2} \) contradict Lemma 2.3.1. This proves that the second bullet holds.

Now suppose that the third bullet does not hold. Since

\[
\chi(N^r(w_{j+1}) \cap N^{r-1}(N(w_{j+1}) \cap P_j)) \leq 3(l + 1)^3 \tau^2,
\]

it follows that

\[
\chi(N^r(w_j) \cap N^{r-1}((N(w_j) \setminus N(w_{j+1})) \cap P_j)) \geq 3(l + 1)^3 \tau^2.
\]
This contradicts Lemma 2.3.1 applied to \( w_j, w_{j+1} \) and the sets \( (N(w_j) \setminus N(w_{j+1})) \cap P_j \) and \( N(w_{j+1}) \cap P_{j+2} \), since the latter two sets are anticomplete to each other. This implies that the third bullet holds.

Now this proves that \( w_1 \in B'_2 \) satisfies the third bullet, contrary to the fact that 
\[
\chi(N^r(u) \cap N^{r-1}(N(u) \cap B'_1)) \geq 3(l + 1)^3 \tau^2
\]
for every \( u \in B'_2 \) by the definition of \( B'_2 \), and \( B'_1 \subseteq P_1 \). This concludes the proof.

\[\square\]

### 2.4 The uncontrolled case

In this section, we deal with the case when \( G \) is not \((r, \varphi)\)-controlled for any small values of \( r \).

Let \( G \) be a graph. A long but short path of length \( k \) in \( G \) is an induced path \( P \) in \( G \) such that the ends of \( P \) are \( u \) and \( v \), and \( d(u, v) = k \) and \( d(u, x) < k \) and \( d(v, x) < k \) for all \( x \in V(P) \setminus \{u, v\} \).

The following lemma proves that long but short paths exist in the graphs we consider.

**Lemma 2.4.1.** Let \( r, \tau \in \mathbb{N} \), and let \( G \) be a graph. Suppose that \( \chi(N^r[v]) \leq \tau \) for all \( v \in V(G) \). Let \( X \subseteq V(G) \) with \( \chi(X) > \tau \). Then for all \( k \leq r + 1 \), there is a long but short path \( P \) of length \( k \) in \( G \) with \( V(P) \subseteq X \).

**Proof.** Let \( G, r, \tau, X \) be as in the statement of the lemma. We may assume that \( G|X \) is connected by replacing \( X \) with the vertex set of a connected component of maximum chromatic number in \( G|X \).

Since \( \chi(X) > \tau \), it follows that for all \( v \in X \), there exists \( u \in X \) such that \( d_G(u, v) > r \). Let \( P \) be a shortest \( u-v \)-path in \( G|X \). Since \( G|X \) is connected, such a path exists. We let \( u = v_1, \ldots, v_l = v \) denote the vertices of \( P \) in order. It follows that \( d(u, v) \geq k \).
While there is a vertex \( v_i \in V(P) \) such that \( d(u, v_i) \geq k \), we replace \( u \) by \( v_i \) and \( P \) by \( P|\{v_1, \ldots, v_i\} \). While there is a vertex \( v_i \in V(P) \) such that \( d(v, v_i) \geq k \), we replace \( v \) by \( v_i \) and \( P \) by \( P|\{v_1, \ldots, v_i\} \). We repeat both of these steps as much as possible. This terminates, since \( P \) becomes shorter at every step. We let \( P' \) denote the path at termination and let \( u', v' \) denote the ends of \( P' \). It follows that \( d(u', v') = k \) (since \( d(u', v') > k \) would imply \( d(v_1, v') \geq k \)). It follows that \( P' \) is a long but short path of length \( k \) in \( G \). \( \square \)

Next we show that a long but short path is helpful for growing long paths starting at one of its neighbors.

**Lemma 2.4.2.** Let \( G \) be a graph. Let \( P \) be a long but short path of length \( r \), and let \( a \in N(V(P)) \). Then there is an induced path \( P' \) of length at least \( r/2 \) in \( G|(V(P) \cup \{a\}) \) such that one of the ends of \( P' \) is \( a \).

**Proof.** Let \( G, P, r, a \) be as in the statement of the lemma. Let \( b \) and \( c \) be the ends of \( P \). Then \( d(b, c) = r \). By the triangle inequality, it follows that either \( d(a, b) \geq r/2 \) or \( d(a, c) \geq r/2 \). By symmetry, we may assume that \( d(a, b) \geq r/2 \). Since \( G|(V(P) \cup \{a\}) \) is connected, we may choose a shortest path \( P' \) from \( a \) to \( b \) in \( G|(V(P) \cup \{a\}) \). It follows that \( P' \) is induced and has length at least \( r/2 \) as claimed. This concludes the proof. \( \square \)

Let \( G \) be a graph and let \( f : V(G) \to \mathbb{N} \) be a coloring of \( G \). Let \( H \) be an induced subgraph of \( G \). Then \( H \) is called \( f\)-rainbow (or just rainbow if \( f \) is clear from context) if \( f(v) \neq f(w) \) for all \( v, w \in V(H) \) with \( v \neq w \).

We need the following result:

**Lemma 2.4.3** (Scott, Seymour \([40]\)). For all \( \kappa, s \in \mathbb{N} \), there exists a \( C = C(s, \kappa) \in \mathbb{N} \) such that for every graph \( G \) with \( \omega(G) \leq \kappa \) and \( \chi(G) > C \), and for every coloring \( f \)
of $G$, there exists an induced path $P$ of $G$ such that $P$ has length at least $s$ and $P$ is $f$-rainbow.

We are ready to prove the main result of this section. Recall that we fixed $T$ and $l \in \mathbb{N}$ such that $T$ is contained in the double star obtained by joining the centers of two $l$-stars by an edge; and in this section, we further assume that $T$ is an H-graph.

**Lemma 2.4.4.** Let $r = l + 2$ and let $\kappa, \tau \in \mathbb{N}$. Then there exists $c$ such that the following holds. Let $G$ be a $T$-free graph with $\omega(G) \leq \kappa$ and such that $\chi(N^r[v]) \leq \tau$ for all $v \in V(G)$, and $\chi(H) \leq \tau$ for every induced subgraph $H$ of $G$ with $\omega(H) < \kappa$. Then $\chi(G) \leq c$.

**Proof.** Let $G, r, \kappa, \tau$ be as in the statement of the lemma. Let $\mathcal{P} = \{P_1, \ldots, P_t\}$ be a collection of long but short paths of length $l$ with the following properties:

- for $i \neq j$, each vertex of $P_i$ has distance at least three from each vertex of $P_j$; and

- subject to satisfying the statement of the first bullet, $|\mathcal{P}|$ is maximal.

Let $V(\mathcal{P}) = V(P_1) \cup \cdots \cup V(P_t)$. Let $S$ be the set of vertices of $G$ at distance at least three from $V(\mathcal{P})$. By Lemma 2.4.1, since $G|S$ does not contain a long but short path of length $l$, it follows that $\chi(G|S) \leq \tau$. Let $A = N(V(\mathcal{P}))$ and $B = N^2(V(\mathcal{P}))$. We first prove that $G|A$ has bounded chromatic number.

\begin{equation}
\text{(2.5) There is a constant } C = C(l) \text{ such that } \chi(A) \leq C\tau.
\end{equation}

To prove this, we proceed as follows. By Lemma 2.4.3, it follows that there is a constant $C = C(2l + 1, \kappa)$ such that for every graph $G$ with $\chi(G) \geq C$ and for every proper coloring $f$ of $G$, there is an induced rainbow path in $G$ of length $\geq 2l + 1$. 

41
Let \( P_i \in \mathcal{P} \), and let \( a \) be an end of \( P_i \). Then \( N(P_i) \subseteq N^{l+1}[a] \subseteq N^r[a] \). This implies that \( \chi(N(P_i)) \leq \tau \). Since the distance between every two vertices in different paths in \( \mathcal{P} \) is at least three, every vertex \( a \in A \) is contained in \( N(P_i) \) for exactly one \( i \in \{1, \ldots, t\} \).

Now let \( g: A \rightarrow \{1, \ldots, \tau\} \) be a function such that \( g|N(P_i) \) is a proper coloring of \( G|N(P_i) \) for all \( i \in \{1, \ldots, k\} \). It follows that there exists a \( j \in \{1, \ldots, \tau\} \) such that \( \chi(g^{-1}(j)) \geq \chi(A)/\tau \). Let \( A' = g^{-1}(j) \), and let \( f: A' \rightarrow \{1, \ldots, t\} \) such that \( f(a) = i \) if and only if \( a \in N(P_i) \).

Suppose that \((2.5)\) does not hold. Then \( \chi(A') > C \), and so by Lemma \[2.4.3\] it follows that \( G|A' \) contains an \( f \)-rainbow path \( P \) of length \( 2l + 1 \). It follows that \( P \) has \( 2l + 2 \) vertices; let \( a, b \) be its two middle vertices (i.e. \( P \setminus \{a, b\} \) consists of two connected components with \( l \) vertices each). Let \( i, j \in \{1, \ldots, t\} \) such that \( a \in N(P_i) \) and \( b \in N(P_j) \).

By Lemma \[2.4.2\] it follows that there are paths \( P_i' \) with \( V(P_i') \subseteq \{a\} \cup V(P_i) \) and \( P_j' \) with \( V(P_j') \subseteq \{b\} \cup V(P_j) \) such that \( a \) is an end of \( P_i' \), \( b \) is an end of \( P_j' \), and both \( P_i', P_j' \) are induced paths of length at least \( l \). But then \( G|(V(P) \cup V(P_i') \cup V(P_j')) \) is an \( H \)-graph that contains \( T \). This is a contradiction, and so \((2.5)\) is proved.

In light of \((2.5)\), to prove the statement of the lemma, it now suffices to prove the following, which we will do in several steps.

\[ \chi(B) \leq C' \chi(A). \] (2.6)

Suppose not. Let \( f: A \rightarrow \mathbb{N} \) be a proper coloring of \( A \) with \( \chi(A) \) colors. Then there exists an \( i \in \{1, \ldots, \chi(A)\} \) such that \( \chi(N(f^{-1}(i)) \cap B) \geq \chi(B)/\chi(A) \). Choose such an \( i \), and let \( A' = N(f^{-1}(i)) \) and \( B' = N(f^{-1}(i)) \cap B \).
For $i \in \{1, \ldots, t\}$, we let $A_i = A' \cap N(P_i)$. For $b \in B'$, we write $A(b) = \{i \in \{1, \ldots, t\} : N(b) \cap A_i \neq \emptyset\}$. We need the following claim:

\begin{equation}
\text{(2.7) For every edge } ab \in E(G|B'), \text{ either } |A(a) \setminus A(b)| < 2 \text{ or } |A(b) \setminus A(a)| < 2. \tag{2.7} \end{equation}

Suppose not. Let $a, b \in B'$ be adjacent such that $|A(a) \setminus A(b)| \geq 2$ and $|A(b) \setminus A(a)| \geq 2$. Let $i, j \in A(a) \setminus A(b)$ and $k, h \in A(b) \setminus A(a)$. By Lemma 2.4.2, there exist paths $P_i', P_j', P_k', P_h'$, each of length at least $l$ and such that $V(P_i') \subseteq \{a\} \cup V(P_i)$ and $V(P_j') \subseteq \{a\} \cup V(P_j)$ and $V(P_k') \subseteq \{b\} \cup V(P_k)$ and $V(P_h') \subseteq \{b\} \cup V(P_h)$ and $a$ is an end of $P_i', P_j'$ and $b$ is an end of $P_k', P_h'$. Now it follows that $G|\left(V(P_i') \cup V(P_j') \cup V(P_k') \cup V(P_h')\right)$ is an H-graph that contains $T$ as an induced subgraph, a contradiction. This proves our claim \text{(2.7)}.

For $i \in \{1, \ldots, 2\kappa^2\}$, we let $B_i = \{b \in B' : |A(b)| = i\}$; and we let $B^* = B' \setminus (B_1 \cup B_2 \cup \cdots \cup B_{2\kappa^2})$. Our claim \text{(2.6)} will follow by bounding $\chi(B_i)$ for $i \in \{1, \ldots, 2\kappa^2\}$ and $\chi(B^*)$ separately, as follows.

\begin{equation}
\chi(B^*) \leq \tau. \tag{2.8} \end{equation}

Suppose not. Since $\chi(H) \leq \tau$ for every induced subgraph $H$ of $G$ with $\omega(H) < \kappa$, it follows $\omega(G|B^*) = \kappa)$. Let $K \subseteq B^*$ be a clique with $|K| = \kappa$. We let $I = \{i \in \{1, \ldots, k\} : i \in \bigcap_{k \in K} A(k)\}$. We first claim that $|I| \leq 3(\kappa-1)\kappa/2$. To prove this, we consider $i \in I$. For $c, d \in K$, we say that $i$ \emph{splits} the edge $cd$ if there is an induced path $a-c-d-b$ with $a, b \in A_i$.

Suppose that there is an edge $cd$ of $G|K$ such that $cd$ is split by at least four elements of $I$, say by $i, j, k, h$. Let $a_i, a_j, a_k, a_h$ such that $a_s \in A_s$ for $s \in \{i, j, k, h\}$,
and such that \( a_i, a_j \in N(c) \setminus N(d) \) and \( a_k, a_h \in N(d) \setminus N(c) \). These vertices exist since \( cd \) splits each of \( i, j, k, h \). Let \( G' \) arise from \( G \) by deleting \( A_s \setminus \{a_s\} \) for \( s \in \{i, j, k, h\} \). Then in \( G' \), we have \( i, j \in A(c) \setminus A(d) \) and \( k, h \in A(d) \setminus A(c) \), contrary to (2.7). It follows that no edge of \( G'|K \) is split by more than three distinct \( s \in I \).

Since \( K \) is a maximum clique of \( G \), it follows that for every vertex \( a \in A_i \), \( N(a) \cap K \neq K \). Let \( a \in A_i \) with \( |N(a) \cap K| \) maximum, and let \( b \in A_i \) such that \( (N(b) \setminus N(a)) \cap K \neq \emptyset \). This is possible by the definition of \( I \), and since \( N(a) \cap K \neq K \). Let \( c \in (N(b) \setminus N(a)) \cap K \). Since \( |N(a) \cap K| \geq |N(b) \cap K| \), it follows that \( (N(a) \setminus N(b)) \cap K \neq \emptyset \); let \( d \in (N(a) \setminus N(b)) \cap K \). Then \( a-d-c-b \) is induced, and so \( i \) splits \( cd \). Since \( K \) contains \((\kappa - 1)\kappa/2\) edges, and since no edge is split by more than three distinct \( s \in I \). It follows that our first claim, that \(|I| \leq 3(\kappa - 1)\kappa/2\), holds.

We now return to the proof of (2.8). Let \( k \in K \) be chosen with \( |A(k)| \) minimum. It follows that \( |A(k) \setminus I| \geq 2\kappa^2 - 3(\kappa - 1)\kappa/2 = \kappa(\kappa/2 + 3/2) > \kappa - 1 \). By the definition of \( I \), it follows that for every \( i \in A(k) \setminus I \), there exists \( k(i) \in K \setminus \{k\} \) such that \( i \notin A(k(i)) \). Since \( |A(k) \setminus I| > |K \setminus \{k\}| \), it follows that there is a \( k' \in K \) such that there exist two distinct \( i, j \in A(k) \setminus I \) with \( k(i) = k(j) = k' \). Since \( i, j \in A(k) \setminus A(k') \) and \( |A(k')| \geq |A(k)| \), it follows that \( |A(k') \setminus A(k)| \geq 2 \) and \( |A(k) \setminus A(k')| \geq 2 \). Since \( k, k' \in K \), it follows that \( kk' \in E(G) \). This contradicts (2.7). This implies that \( \omega(B^*) < \tau \), which proves (2.8).

There is a constant \( C = C(l, \tau) \) such that \( \chi(B_i) \leq C \) for all \( i \in \{1, \ldots, 2\kappa^2\} \).

(2.9)

Suppose not; and let \( i \in \{1, \ldots, 2\kappa^2\} \) be such that (2.9) does not hold for \( i \). For \( b \in B_i \), we write \( A(b) = \{A(b)^1, \ldots, A(b)^l\} \) such that \( A(b)^1 \leq A(b)^2 \leq \cdots \leq A(b)^l \).
We claim that for every \( j \in \{0, \ldots, i\} \), there exists \( B_j^i \subseteq B_i \) such that for every edge \( ab \) with \( a, b \in B_j^i \), we have that \( A(a)^k \neq A(b)^k \) for all \( k \in \{1, \ldots, j\} \), and \( \chi(B_j^i) \geq \chi(B_i)/\tau^j \). We prove this claim by induction; it holds for \( j = 0 \) by letting \( B_i^0 = B_i \).

Now suppose that the statement for some \( j < i \) and let \( B_j^i \) be the corresponding set; we will prove that it is true for \( j + 1 \) as well. Since \( N^2(P_k) \subseteq N^\tau(a) \), where \( a \) is an end of \( P_k \), it follows that for every \( k \in \{1, \ldots, t\} \), \( \chi(N^2(P_k)) \leq \tau \). Let \( B_i^j(k) = \{b \in B_i^j : A(b)^j = k\} \) for \( k \in \{1, \ldots, t\} \). It follows that \( B_i^j(k) \subseteq N^2(P_k) \) and therefore, \( \chi(B_i^j(k)) \leq \tau \). Moreover, \( \{B_i^j(k)\}_{k \in \{1, \ldots, t\}} \) is a partition of \( B_i^j \). Let \( f : B_i^j \to \{1, \ldots, \tau\} \) be such that \( f|B_i^j(k) \) is a proper coloring of \( B_i^j(k) \) for all \( k \in \{1, \ldots, t\} \). Then there exists a \( h \in \{1, \ldots, t\} \) such that \( \chi(f^{-1}(h)) \geq \chi(B_i^j)/\tau \).

We let \( B_i^{j+1} = f^{-1}(h) \). Since \( B_i^j(k) \cap f^{-1}(h) \) is stable for all \( k \in \{1, \ldots, t\} \), it follows that \( B_i^{j+1} \) has the desired properties. This proves our claim.

By the claim above, there exists \( D \subseteq B_i \) such that for every edge \( ab \) with \( a, b \in B_j^i \), we have that \( A(a)^k \neq A(b)^k \) for all \( k \in \{1, \ldots, i\} \), and \( D \) satisfies that \( \chi(D) \geq \chi(B_i)/\tau^i \). We now proceed with the proof of (2.9); we need the following statement:

\[
\text{(2.10)} \quad \text{Let } ab \in E(G|D). \text{ Then } |A(a) \setminus A(b)| = 1, \text{ and either } A(a)^j = A(b)^{j+1} \text{ for } j \in \{1, \ldots, i-1\} \text{ or } A(b)^j = A(a)^{j+1} \text{ for } j \in \{1, \ldots, i-1\}. 
\]

The first statement \( |A(a) \setminus A(b)| = 1 \) follows from (2.7) and since \( A(a) \neq A(b) \) by the choice of \( D \).

We may assume that \( A(a)^1 < A(b)^1 \) by symmetry and by the choice of \( D \). Thus \( A(a)^1 \notin A(b) \), and consequently \( A(a) \setminus \{A(a)^1\} \subseteq A(b) \). Since \( A(a)^j \neq A(b)^j \) for all \( j \in \{1, \ldots, i\} \), it follows that \( A(b)^j = A(a)^{j+1} \) for all \( j \in \{1, \ldots, i-1\} \), and thus (2.10) is proved.

\[
\text{(2.11)} \quad \text{If } i \geq 2, \text{ then } G|D \text{ is triangle-free.}
\]
Suppose not, and let \( \{a, b, c\} \) be pairwise adjacent. Without loss of generality, we may assume that \( A(a)^1 < A(b)^1 < A(c)^1 \). Then, by (2.10), it follows that \( A(a)^2 = A(b)^1 \) and \( A(a)^2 = A(c)^1 \), and therefore \( A(b)^1 = A(c)^1 \), contrary to the choice of \( D \). It follows that \( G|D \) is triangle-free, and (2.11) is proved.

If \( \kappa \geq 3 \), this implies that \( \chi(D) \leq \tau \) and thus \( \chi(B_i) \leq \chi(D) \cdot \tau^i \leq \tau^{i+1} \), and (2.9) follows. Therefore, we may assume in the following that either \( \kappa = 2 \), i.e. \( G \) is triangle-free, or \( i = 1 \).

There is a constant \( C = C(l, \tau) \) such that \( \chi(B_i) \leq C \) for all \( i \in \{3, \ldots, 2\kappa^2\} \).

Suppose not; and let \( i \in \{3, \ldots, 2\kappa^2\} \) such that (2.12) does not hold for \( i \). If \( G|D \) is stable, then \( \chi(B_i) \leq \tau^i \), so we may assume that there is an edge \( ab \) with \( a, b \in D \).

We may assume that \( A(a)^1 < A(b)^1 \). Since \( i \geq 3 \), it follows that \( A(a)^1 < A(a)^2 < A(b)^2 < A(b)^3 \). Let \( a_j \in N(a) \cap N(P_{A(a)^j}) \) for \( j \in \{1, 2\} \) and \( b_j \in N(b) \cap N(P_{A(b)^j+1}) \) for \( j \in \{1, 2\} \).

Since \( i > 1 \), it follows that \( G \) is triangle-free, and consequently, \( b \) is non-adjacent to \( a_1, a_2 \) and \( a \) is non-adjacent to \( b_1, b_2 \). By Lemma 2.4.2, there are paths \( P(a_1), P(a_2), P(b_1), P(b_2) \) of length at least \( l \) with \( P(a_j) \subseteq \{a_j\} \cup V(P_{A(a)^j}) \) and such that \( P(a_j) \) starts at \( a_j \) for \( j \in \{1, 2\} \), and \( P(b_j) \subseteq \{b_j\} \cup V(P_{A(b)^j+1}) \) and such that \( P(b_j) \) starts at \( b_j \) for \( j \in \{1, 2\} \). But then \( G|((\{a, b\} \cup V(P(a_1)) \cup V(P(a_2)) \cup V(P(b_1)) \cup V(P(b_2))) \) is an H-graph that contains \( T \), a contradiction. This proves (2.12).

There is a constant \( C = C(l, \tau) \) such that \( \chi(B_i) \leq C \) for all \( i \in \{1, 2\} \).
To prove (2.13), we proceed as follows. By Lemma 2.4.3 it follows that there is a constant $C = C(4t + 4, \kappa)$ such that for every graph $G$ with $\chi(G) \geq C$ and for every proper coloring $f$ of $G$, there is an induced rainbow path in $G$ of length at least $3l + 2$.

We let $f : D \to \{1, \ldots, t\}$ be a function with $f(a) = A(a)^1$. We may assume that

$$\chi(D) \geq C,$$

for otherwise $\chi(B_i) \leq \tau^i \cdot C$ and (2.13) follows. Therefore, it follows that $G|D$ contains a rainbow path $P'$ of length $4l + 4$. Let $v_1, \ldots, v_{4l+5}$ denote the vertices of $P'$ in order.

We first assume that $i = 2$. We claim that if $i = 2$, then there is no $k \in \{2, 3, \ldots, 4l + 4\}$ such that $A(v_k)^1 > A(v_{k+1})^1 < A(v_{k+2})^1$. Suppose that there is such a $k$; then (2.10) implies that $A(v_k)^1 = A(v_{k+1})^2 = A(v_{k+2})^1$, contrary to $P'$ being rainbow. It follows that there exists a $k \in \{1, \ldots, 4l + 5\}$ such that $A(v_1)^1, \ldots, A(v_k)^1$ is a strictly increasing sequence and $A(v_k)^1, \ldots, A(v_{4l+5})^1$ is a strictly decreasing sequence. Therefore, there is a subpath $P''$ of $P'$ such that $P''$ has length at least $2l + 1$, and such that $w_1, \ldots, w_{2l+2}$ are the vertices of $P''$ in order, and $A(w_1)^1, \ldots, A(w_{2l+2})^1$ is either a strictly increasing or a strictly decreasing sequence.

We let $P''$ be the path defined above if $i = 2$, and we let $P''$ be an arbitrary subpath of $P'$ of length $2l + 1$ if $i = 1$.

Let $a, b$ be the two middle vertices of $P''$, i.e. $a, b \in V(P')$ such that $P'' \setminus \{a, b\}$ consists of two connected components with $l$ vertices each. By symmetry, we may assume that $A(a)^1 < A(b)^1$. Let $c \in N(a) \cap N(P_{A(a)^1})$ and $d \in N(b) \cap N(P_{A(b)^1})$. If $i = 1$, then since $P'$ is rainbow, it follows that $N(c) \cap V(P'') = \{a\}$ and $N(d) \cap V(P'') = b$. If $i = 2$, we let $w_1, \ldots, w_{2l+2}$ denote the vertices of $P''$ in order such that $a = w_{l+1}$. Then $A(w_1)^1, \ldots, A(w_{2l+2})^1$ is strictly increasing by the claim above. It follows that $A(w_k)^1 < A(w_{k+1})^1 < A(w_{k+2})^1$ for all $k \leq l - 1$, and so $c, d$ are non-adjacent to $w_1, \ldots, w_{l-1}$, and $d$ is non-adjacent to $w_l$. Since $A(w_k)^2 > A(w_k)^1 > A(w_{k+2})^1$ for $k \geq l + 3$, it follows that $c, d$ are non-adjacent to $w_{l+3}, \ldots, w_{2l+2}$. Since
\(i = 2\), it follows that \(G\) is triangle-free, and hence \(c\) is non-adjacent to \(w_l\) and \(w_{l+2}\) and \(d\) is non-adjacent to \(w_{l+1}\). This implies that \(N(c) \cap V(P'') = \{a\}\) and \(N(d) \cap V(P'') = b\). Therefore, we have that \(N(c) \cap V(P'') = \{a\}\) and \(N(d) \cap V(P'') = b\) in both case \((i = 1\) and \(i = 2)\).

Now there exists a path \(P_1^*\) of length at least \(l\) starting at \(c\) and otherwise contained in \(V(P_{A(a)})\) and a path \(P_2^*\) of length at least \(l\) starting at \(d\) and otherwise contained in \(V(P_{A(b)})\). It follows that \(G|(V(P_1^*) \cup V(P_2^*) \cup V(P''))\) is an \(H\)-graph containing \(T\), a contradiction. This proves \((2.13)\).

From \((2.12)\) and \((2.13)\), it follows that \((2.9)\) holds.

Together, \((2.8)\), \((2.9)\) imply that there is a constant \(C = C(l, \tau)\) such that \(\chi(B') \leq \tau + 2\kappa^2 C(l, \tau)\), and thus \(\chi(B) \leq \chi(A) \cdot (\tau + 2\kappa^2 C(l, \tau))\). This proves \((2.6)\).

By combining \((2.5)\) and \((2.6)\), it follows that \(\chi(G) \leq C \cdot C'\), and thus the result is proved.

We are now ready to prove the main result of this chapter.

**Proof of Theorem 2.0.3.** Let \(T\) be an \(H\)-graph that is contained in a double star obtained by joining the centers of two disjoint \(l\)-stars. We proceed by induction on the clique number \(\kappa\). If \(\kappa = 1\), then every \(T\)-free graph \(G\) with \(\omega(G) \leq \kappa\) satisfies \(\chi(G) \leq 1\). Therefore, we may assume that \(\kappa > 1\) and there is a \(\tau > 0\) such that every \(T\)-free graph \(G\) with \(\omega(G) < \kappa\) satisfies \(\chi(G) < \tau\).

By Lemma \(2.4.4\), we know that for every \(\kappa\) and \(\tau' \geq \tau\), there exists a constant \(c\) such that every graph \(G\) with \(\chi(N^{l+2}[v]) \leq \tau'\) for all \(v \in V(G)\) satisfies \(\chi(G) \leq c\). This implies that there is a non-decreasing function \(\varphi\) such that for every \(T\)-free graph \(G\) with \(\omega(G) \leq \kappa\), there is a vertex \(v \in V(G)\) such that \(\chi(G) \leq \varphi(\chi(N^{l+2}_G[v]))\). This implies that every \(T\)-free graph \(G\) with \(\omega(G) \leq \kappa\) is \((l + 2, \varphi)\)-controlled.
Since $\chi(N(v)) \leq \tau$ for all $v \in V(G)$, it follows that if there is a function $\varphi$ such that every $T$-free graph $G$ is $(1, \varphi)$-controlled, then $\chi(G) \leq \varphi(\tau)$ for all $T$-free graphs $G$, and Theorem 2.0.3 is proved. Thus we may assume that there is a minimum $r$ such that there is a function $\varphi$ such that every $T$-free graph $G$ with $\omega(G) \leq \kappa$ is $(r, \varphi)$-controlled, but there exists a $\tau'$ such that for every $\tau''$, there is a $T$-free graph $G$ with $\omega(G) \leq \kappa$, $\chi(G) \geq \tau''$ and $\chi(N^{r-1}[v]) \leq \tau'$ for all $v \in V(G)$. We may assume that $\tau' \geq \tau$.

We now apply Lemma (2.3.4). For every function $\psi$ there exists a constant $c$ such that for every $G$ that satisfies the conditions of Lemma 2.3.4 with $\psi$, we have that $\chi(G) \leq c$. Supposing that Theorem 2.0.3 does not hod, this implies that there exists a $\gamma$ such that for every $\gamma'$, there is a $T$-free $(r, \varphi)$-controlled graph $G$ such that $\chi(G) \geq \gamma'$, but $\chi(N^{r-1}[v]) \leq \tau'$ for all $v \in V(G)$ and $\chi(L_r) \leq \gamma$ for every $\tau$-conservative $r$-leveling with an $l$-spire $L = (v_1, \{v_2\}, \ldots, \{v_l\}, \{v\}, L_1, \ldots, L_r)$.

By Lemma 2.3.3 this implies that for every $\gamma'$, there is a $T$-free $(r, \varphi)$-controlled graph $G$ such that $\chi(N^{r-1}[v]) \leq \tau'$ and $\chi(G) \geq \gamma'$, but $\chi(L_r) \leq (\gamma + (2l+1)\tau)\tau$ for every $\tau$-conservative $r$-leveling $L = (\{v\}, L_1, \ldots, L_r)$ in $G$. But this contradicts Lemma 2.3.6 which proves Theorem 2.0.3.\qed
Chapter 3

Complete or anticomplete pairs

An \((x, y)\)-pair in a graph \(G\) is a pair \(A, B \subseteq V(G)\) of disjoint sets such that \(|A| \geq x\) and \(|B| \geq y\). An \((x, y)\)-pair \(A, B\) is

- complete, if \(A\) is complete to \(B\);
- anticomplete, if \(A\) is anticomplete to \(B\);
- \(c\)-sparse for \(c \geq 0\) if \(|E(A, B)| \leq c|A||B|\); and
- \(c\)-dense for \(c \geq 0\) if \(|E(A, B)| \geq c|A||B|\).

Let \(\varepsilon > 0\). A graph \(G\) is \(\varepsilon\)-sparse if \(|N[v]| < \varepsilon|V(G)|\) for all \(v \in V(G)\), and \(\varepsilon\)-dense if \(G^c\) is \(\varepsilon\)-sparse. We begin with a simple observation:

**Lemma 3.0.1.** Let \(G\) be an \(n\)-vertex graph, and let \(y \geq x > 0\) such that \(G\) has no anticomplete \((x, y)\)-pair. Then \(G\) has a vertex of degree at least \((n - \lceil y \rceil + 1)/\lceil x \rceil - 1\).

**Proof.** We may assume that \(x + y \leq n\). Let \(X\) be a set of \(\lceil x \rceil\) vertices. It follows that \(X\) has at most \(\lfloor y \rfloor - 1\) non-neighbors in \(G\). Consequently, there is a vertex in \(X\) with degree at least \((n - \lfloor y \rfloor - \lceil x \rceil + 1)/\lceil x \rceil = (n - \lfloor y \rfloor + 1)/\lceil x \rceil - 1\). \(\square\)
3.1 The Erdős-Hajnal conjecture

The Erdős-Hajnal conjecture is one of the most well-known open conjectures related to the study of graphs with forbidden induced subgraphs:

**Conjecture 3.1.1** (Erdős, Hajnal [21, 22]). For every graph \( H \), there is a constant \( c > 0 \) such that every \( H \)-free graph \( G \) contains a clique or a stable set of size at least \( |V(G)|^c \).

We say that \( H \) satisfies the Erdős-Hajnal conjecture if the statement of Conjecture 3.1.1 holds for \( H \). We say that a class \( C \) of graphs has the **EH-property** if there is a constant \( c > 0 \) such that every graph \( G \in C \) contains a clique or a stable set of size at least \( |V(G)|^c \). The following result implies that the class of all graphs does not have the EH-property:

**Theorem 3.1.2** (Erdős [19]). There is a constant \( c > 0 \) such that for every sufficiently large \( n \), there is an \( n \)-vertex graph \( G \) with \( \alpha(G), \omega(G) \leq c \log_2 n \).

The Erdős-Hajnal conjecture states that every hereditary proper subclass of the class of all graphs has the EH-property. This conjecture is known to be true for some graphs \( H \), including the bull [12], a graph obtained from a four-vertex path \( a-b-c-d \) by adding a new vertex \( e \) joined to \( b \) and \( c \) and non-adjacent to \( a \) and \( d \); and all induced subgraphs of the bull. The following theorem is currently the only known construction for building new graphs \( H \) that satisfy the Erdős-Hajnal conjecture from other graphs that do:

**Theorem 3.1.3** (Alon, Pach, Solymosi [2]). Let \( H, H' \) be graphs, and let \( v \in V(H) \). Let \( H'' \) be the graph obtained by taking the disjoint union of \( H \setminus \{v\} \) and \( H' \), and making the vertices of \( H' \) complete to \( N_H(v) \). If \( H, H' \) satisfy the Erdős-Hajnal conjecture, then so does \( H'' \).
Theorem 3.1.3 together with [12], implies that Conjecture 3.1.1 holds for all four-vertex graphs, and for all five-vertex graphs except for \( C_5, P_5 \) and \( P_5^c \) [8]. On the other hand, the conjecture is open for the three graphs mentioned [8].

Finding sparse, dense, complete, or anticomplete pairs in \( H \)-free graphs (or \( \{H, H^c\} \)-free graphs) is a common strategy for proving variants of the Erdős-Hajnal conjecture. The motivation for this is the following. A graph is a cograph if it is \( P_4 \)-free. A graph \( G \) is perfect if \( \chi(H) = \omega(H) \) for every induced subgraph \( H \) of \( G \).

**Theorem 3.1.4** (Chudnovksy, Robertson, Seymour, Thomas [11]). A graph is perfect if and only if it contains no induced subgraph isomorphic to an odd hole or odd antihole.

Since \( \alpha(G) \geq |V(G)|/\chi(G) \) in general (because a coloring is a partition of \( V(G) \) into \( \chi(G) \) parts, each of size at most \( \alpha(G) \)), it follows that \( \alpha(G) \omega(G) \geq |V(G)| \) for every perfect graph. By Theorem 3.1.4 and since every odd hole and every odd antihole contains \( P_4 \), it follows that cographs are perfect (which was also proved in [41]). This motivates two equivalent definitions of the EH-property [8]:

- A class \( C \) of graphs has the EH-property if and only if there is a constant \( c > 0 \) such that every graph \( G \in C \) satisfies \( \alpha(G) \cdot \omega(G) \geq |V(G)|^c \).

- A class \( C \) of graphs has the EH-property if and only if there is a constant \( c > 0 \) such that every graph \( G \in C \) contains a \( P_4 \)-free induced subgraph of size at least \( |V(G)|^c \).

Since \( P_4 \) is connected and isomorphic to \( P_4^c \), it follows that if \( G \) and \( G' \) are \( P_4 \)-free, then both their disjoint union and their join are \( P_4 \)-free as well. Therefore, finding pairs \( A, B \) of big size and with very low or very high edge density between them is a helpful step towards for finding a \( P_4 \)-free induced subgraph of a given size in \( A \), and finding a \( P_4 \)-free induced subgraph of a given size in \( B \), that either have no edges between them or such that all edges between them are present, thus yielding
a bigger $P_4$-free induced subgraph. When $A$ and $B$ have linear size and the edge density between them is 0 or 1, this is sufficient for proving the EH-property; see Section 3.1.1.

As an approximate solution to Conjecture 3.1.1, the following question is of interest:

**Question 3.1.5.** For which functions $f(n, c)$ is it true that for every graph $H$, there exists a constant $c$ such that every $H$-free graph $G$ contains a clique or a stable set of size at least $f(n, c)$, where $n = |V(G)|$?

Solving Conjecture 3.1.1 is equivalent to solving the above question for the function $f(n, c) = n^c$. The following result implies that Question 3.1.5 is true for $f(n, c) = 2^{c\sqrt{\log_2 n}}$:

**Theorem 3.1.6** (Erdős, Hajnal [22]). For every graph $H$, there is a constant $c > 0$ such that every $H$-free graph $G$ contains a clique or a stable set of size at least $2^{c\sqrt{\log_2 n}}$, where $n = |V(G)|$.

Another approximate solution for Conjecture 3.1.1 is given by the following result, which shows that the Erdős-Hajnal conjecture holds for almost all graphs:

**Theorem 3.1.7** (Loebl, Reed, Scott, Thomason, Thomassé [35]). For every graph $H$, there is a constant $c > 0$ such that as $n \to \infty$, the probability that a random $H$-free graph on $n$ vertices has a clique or a stable set of size at least $n^c$ converges to 1.

### 3.1.1 Sparse symmetric version

A tool for proving variants of the Erdős-Hajnal conjecture is the following notion: A class $C$ of graphs has the strong EH-property if there is an $\varepsilon > 0$ such that every graph $G \in C$ with $|V(G)| > 1$ contains a complete or anticomplete $(\varepsilon n, \varepsilon n)$-pair.

This property, as the name suggests, is stronger than the EH-property:
Lemma 3.1.8 (Alon, Pach, Pinchasi, Radoićić, Sharir [1], see also [26]). Let \( \mathcal{C} \) be a class of graphs that has the strong EH-property. Then \( \mathcal{C} \) has the EH-property.

The following result of Rödl implies that when considering \( H \)-free graphs \( G \), we can often assume that \( G \) is either sparse or dense:

Theorem 3.1.9 (Rödl [38]). For every graph \( H \) and every \( \varepsilon > 0 \) there is a \( \delta = \delta(H, \varepsilon) > 0 \) such that every \( H \)-free graph \( G \) contains an induced subgraph \( J \) with \( |V(J)| \geq \delta|V(G)| \) and such that \( |E(J)| \leq \varepsilon|V(J)|(|V(J)| - 1)/2 \) or \( |E(J^c)| \leq \varepsilon|V(J)|(|V(J)| - 1)/2 \).

The bounds on \( \delta \) have been improved by Fox and Sudakov [27].

We will use the following version of this result:

Corollary 3.1.10. For every graph \( H \) and every \( \varepsilon > 0 \) there is a \( \delta' > 0 \) such that every \( H \)-free graph \( G \) either satisfies \( |V(G)| \leq 1/\delta' \) or contains an induced subgraph \( J \) with \( |V(J)| \geq \delta'|V(G)| \) and such that either \( J \) or \( J^c \) is \( \varepsilon \)-sparse.

Proof. Let \( H \) be a graph, and let \( \varepsilon > 0 \). Let \( \varepsilon' = \varepsilon/8 \), and let \( \delta = \delta(H, \varepsilon/8) \) as in Theorem 3.1.9. We claim that we may choose \( \delta' = \varepsilon'/8 \). Now let \( G \) be \( H \)-free, and let \( J \) as in Theorem 3.1.9.

By symmetry, we may assume that \( |E(J)| \leq \varepsilon'|V(J)|(|V(J)| - 1)/2 \). If \( |V(J)| \leq 8/\varepsilon' \), then \( |V(G)| \leq |V(J)|/\delta \leq 8/(\varepsilon'\delta) \), and the result holds. Therefore we may assume that \( |V(J)| > 8/\varepsilon' \).

Let \( X \subseteq V(J) \) be the set of vertices \( v \in V(J) \) with \( d(v) \geq 4\varepsilon'|V(J)| - 1 \). Since \( |E(J)| = \frac{1}{2} \sum_{x \in X} d_J(x) \), it follows that

\[
|X| \leq \frac{\varepsilon'|V(J)|(|V(J)| - 1)}{4\varepsilon'|V(J)| - 1}
\leq \frac{(|V(J)| - 1)/4}{4\varepsilon'|V(J)| - 1}
\leq \frac{(|V(J)| - 1)/4}{2 - 1} \leq (|V(J)| - 1)/2,
\]

54
and so $J' = J \setminus X$ satisfies $|V(J')| \geq |V(J)|/2 \geq \delta' |V(G)|$, and for every vertex $v \in V(J')$, we have that $d(v) \leq 4\varepsilon'|V(J)| - 1$, and so $|N_{J'}(v)| \leq 4\varepsilon'|V(J)| \leq 8\varepsilon'|V(J')| = \varepsilon|V(J')|$, which implies that $J$ is $\varepsilon$-sparse, as desired.

This motivates the following definition: A class $\mathcal{C}$ of graphs has the strong sparse \textit{EH-property} if every $\varepsilon$-sparse graph $G \in \mathcal{C}$ with $|V(G)| > 1$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$. We say that a graph $G$ is $\varepsilon$-coherent if

- $G$ is $\varepsilon$-sparse; and

- $G$ contains no anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$.

With this definition, it follows that $\mathcal{C}$ has the strong sparse EH-property if and only if there exists an $\varepsilon > 0$ such that no graph in $\mathcal{C}$ is $\varepsilon$-coherent. Graph classes with this property will be discussed in Section 3.2. This property is useful because of the following lemma, which is an easy consequence of Corollary 3.1.10:

\textbf{Lemma 3.1.11} (Chudnovsky, Scott, Seymour, Spirkl [16]). If $H, H'$ are graphs such that both the class of $H$-free graphs and the class of $H'$-free graphs have the strong sparse EH-property, then the class of $\{H, H^c\}$-free graphs has the strong EH-property.

\textit{Proof.} Let $H, H'$ be as in the statement of the lemma. Let $\varepsilon > 0$ be such that every $H$-free graph and every $H'$-free graph $G$ which is $\varepsilon$-sparse contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair. Let $\delta > 0$ be as in Corollary 3.1.10, i.e. such that every $H$-free graph $G$ with $|V(G)| > 1/\delta$ contains an induced subgraph $J$ with $|V(J)| \geq \delta |V(G)|$ and such that $J$ or $J^c$ is $\varepsilon$-sparse.

Now let $G$ be $H$-free and $H^c$-free, and let $n = |V(G)|$. If $|V(G)| \leq 1/\delta$, then either $|V(G)| = 1$, or $|V(G)|$ contains two vertices $u, v$ with $u \neq v$, and $\{u\}, \{v\}$ is a complete or anticomplete $(\delta n, \delta n)$-pair. Therefore, we assume that $|V(G)| > 1/\delta$ from now on.

By Corollary 3.1.10, it follows that $G$ contains an induced subgraph $J$ with $|V(J)| \geq \delta |V(G)|$ and such that $J$ or $J^c$ is $\varepsilon$-sparse. If $J$ is $\varepsilon$-sparse, then since
\(J\) is \(H\)-free and \(H\)-free graphs have the strong sparse EH-property, it follows that \(J\) contains an anticomplete \((\varepsilon|V(J)|, \varepsilon|V(J)|)\)-pair, which is an anticomplete \((\varepsilon \delta n, \varepsilon \delta n)\)-pair in \(G\). Therefore, we may assume that \(J^c\) is \(\varepsilon\)-sparse. Since \(J^c\) is \(H'\)-free and \(H'\)-free graphs have the strong sparse EH-property, it follows that \(J^c\) has an anticomplete \((\varepsilon|V(J)|, \varepsilon|V(J)|)\)-pair, which is a complete \((\varepsilon \delta n, \varepsilon \delta n)\)-pair in \(G\).

This implies that every \(\{H, H^c\}\)-free graph \(G\) with \(n = |V(G)|\) has a complete or anticomplete \((\varepsilon \delta n, \varepsilon \delta n)\)-pair. This concludes the proof.

In particular, if \(H\)-free graphs have the strong sparse EH-property, then the following weaker version of Conjecture 3.1.1 holds for \(H\):

**Conjecture 3.1.12** (Chudnovsky [8]). For every graph \(H\), there is a constant \(c > 0\) such that every \(\{H, H^c\}\)-free graph \(G\) contains a clique or a stable set of size at least \(|V(G)|^c\).

This conjecture holds for every graph \(H\) which satisfies Conjecture 3.1.1, but it was also proved for some graphs for which Conjecture 3.1.1 is not known to be true. In particular, Conjecture 3.1.12 was proved for the following graphs:

- \(H = P_k\) or \(H = P_k^c\) for \(k \in \mathbb{N}\) [4]
- hooks and their complements [7], where a hook arises from a path \(v_1-v_2-v_3-\ldots-v_k\) by adding a vertex \(w\) adjacent only to \(v_3\).

In Section 3.2, we will prove results about the strong EH-property. In Section 3.3, we consider a weakening, in which we ask for a polynomial-size set and a linear-size set with no edges between them. In Section 3.4, we allow both sets to have polynomial size. Section 3.5 contains improvements for results from Sections 3.3 and 3.4. In Section 3.6, we consider upper bounds on the sizes of the complete or anticomplete pairs. Finally, in Section 3.7, we turn our attention to the case \(H = K_3\), which is the smallest graph for which our methods in Sections 3.3 and 3.5 currently do not work.
### 3.2 Linear anticomplete pairs in sparse graphs

In this section, we consider the following two questions:

**Question 3.2.1.** For which graphs $H$ does there exist an $\varepsilon > 0$ such that every $H$-free graph $G$ contains a complete or anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)| > 1$?

**Question 3.2.2.** For which graphs $H$ does there exist an $\varepsilon > 0$ such that every $\{H, H^c\}$-free graph $G$ contains a complete or anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)| > 1$?

As we will show in Section 3.2.1, a necessary condition for Question 3.2.1 is that $H$ and $H^c$ have to be forests, and hence a necessary condition is that $H$ is an induced subgraph of the four-vertex path $P_4$. Since $P_4$ is self-complimentary, Lemma 3.2.8 implies that $P_4$ has the property in Question 3.2.1 and hence so do all induced subgraphs of $P_4$. Therefore, Lemma 3.2.8 provides a complete answer to Question 3.2.1. For Question 3.2.2, a necessary condition is that $H$ or $H^c$ is a forest; and the following conjecture asserts that this condition is sufficient as well:

**Conjecture 3.2.3** (Liebenau, Pilipczuk [33]). Let $H$ be a forest. Then there is a constant $\varepsilon > 0$ such that every $\{H, H^c\}$-free graph $G$ contains a complete or anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$.

In Section 3.2.3, we will prove this conjecture for a class of trees called caterpillars, and in Section 3.2.5, we will extend the result to all trees, which proves Conjecture 3.2.3.

In Section 3.2.4, we use our results from Section 3.2.3 to prove the conjecture for subdivided caterpillars and consider a further relaxation of the problem in Question 3.2.2.

**Question 3.2.4.** For which graphs $H$ does there exist an $\varepsilon > 0$ such that for every graph $G$, if neither $G$ nor $G^c$ contains an induced subgraph isomorphic to a subdivision...
of $H$, then $G$ contains a complete or anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)| > 1$?

As we show in Theorem 3.2.18, this is true for every graph $H$.

In Section 3.2.6 we consider Question 3.2.2 in directed graphs. In Section 3.2.7 we show that the class of graphs that do not contain a hole of length congruent to $j$ modulo $k$ has the strong sparse EH-property if and only if $j$ is even or $k$ is odd.

### 3.2.1 Necessary conditions

The construction used to prove the following result is due to Erdős [20], who used it to prove that there are graphs of large girth and large chromatic number. Constructions of this kind, bounding the number of cycles in a random graph, are currently our main tool for proving necessary conditions. The implications of Theorem 3.2.5 for the strong EH-property were observed in [7].

**Theorem 3.2.5** (Erdős [20]; Choromanski, Falik, Liebenau, Patel, Pilipczuk [7]). *Let $H$ be a graph and not a forest. Then, for every $\varepsilon > 0$, and for every sufficiently large $n \in \mathbb{N}$, there is an $n$-vertex graph $G$ such that $G$ is $H$-free and $G$ has no anticomplete $(\varepsilon n, \varepsilon n)$-pair. Furthermore, $G$ can be chosen as an $\varepsilon$-sparse graph.*

**Proof.** Let $p = f(n)$, with $f$ to be chosen later, such that $p \geq 1/n$. Let $G \sim G(2n, p)$ be a random graph with $2n$ vertices, and such that every edge is present independently with probability $p$. Then the expected number of cycles of length at most $k = |V(H)|$ in $G$ is bounded by

$$\sum_{j=3}^{k} n^j p^j \leq k(2n)^k p^k.$$

It follows that, if $f(n) \leq n^{1/k-1}/4k$, then the expected number of cycles of length at most $k$ is bounded by $n/2$. It follows that the probability that $G$ contains at least $n$ such cycles is bounded by $1/2$.  

58
Now we consider the probability that $G$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair. There are at most $3^{2n}$ different $(\varepsilon n, \varepsilon n)$-pairs in $G$ (counted by labeling each vertex with one of the three options “in $A$”, “in $B$”, or “neither”), and for a fixed $(\varepsilon n, \varepsilon n)$-pair in $G$, the probability that it is anticomplete in $G$ is at most $(1 - p)^{\varepsilon^2 n^2}$. This implies that the expected number of anticomplete $(\varepsilon n, \varepsilon n)$-pairs in $G$ is bounded from above by 

$$3^{2n} \cdot (1 - p)^{\varepsilon^2 n^2} \leq e^{2n \ln 3} e^{-p\varepsilon^2 n^2}$$

using that $1 - p \leq e^{-p}$. Thus, if $f(n) \geq (1 + 2 \ln 3)/(\varepsilon^2 n)$, this quantity is at most $e^{-n}$, which converges to zero for $n \to \infty$. It follows that, if

$$\frac{1 + 2 \ln 3}{\varepsilon^2} \cdot n^{-1} \leq f(n) \leq n^{1/k - 1}/4k,$$

and if $n$ is sufficiently large, then there is a $2n$-vertex graph $G$ with no anticomplete $(\varepsilon n, \varepsilon n)$-pair and such that $G$ contains at most $n$ cycles of length at most $k$. Let $G'$ arise from $G$ by deleting one vertex from each cycle of length at most $k$. It follows that $G'$ has at least $n$ vertices and no anticomplete $(\varepsilon n, \varepsilon n)$-pair, as desired. This proves the first statement.

For the proof of the second statement, we let $f(n) = \frac{1+2\ln 3}{\varepsilon^2} \cdot n^{-1}$, and let $t = \frac{1+2\ln 3}{\varepsilon^2}$. Now the expected number of cycles of length at most $k$ in $G$ is bounded by $k(2t)^k$, and therefore, the probability that there are at most $4k(2t)^k$ cycles of length at most $k$ in $G$ is bounded by $1/4$. The expected degree of a vertex is $2t$, and so the expected number of vertices with degree at least $64t$ is bounded by $n/16$. Thus, the probability that there are at least $n/4$ such vertices is at most $1/4$. By the union bound, the probability that there is a graph with at most $4k(2t)^k$ cycles of length at most $k$, at most $n/4$ vertices of degree at least $64t$, and no $(\varepsilon n, \varepsilon n)$-anticomplete pair is at least $1 - 1/2 - e^{-n}$, which is positive for $n$ sufficiently large. Therefore, for $n$ sufficiently large, there is a graph, say $G$, such that:
• $G$ contains at most $4k(2t)^k$ cycles of length at most $k$;

• $G$ has no $(\varepsilon n, \varepsilon n)$-anticomplete pair; and

• $G$ has at most $n/4$ vertices of degree at least $64t$.

Now we delete from $G$ a vertex from every cycle of length at most $k$, and every vertex of degree at least $64t$. Let $G'$ denote the resulting graph. It follows that $G'$ is $H$-free, has no $(\varepsilon n, \varepsilon n)$-anticomplete pair, has at least $n$ vertices (if $n$ is sufficiently large), and has maximum degree $64t = 64 \cdot \frac{1+2\ln 3}{\varepsilon^2}$. This implies that if $n$ is sufficiently large, $G'$ is $\varepsilon$-sparse.

Theorem 3.2.5 implies that if $F$ is finite, then the class of $F$-free graphs does not have the strong sparse EH-property unless $F$ contains a forest, and the class of $F$-free graphs does not have the strong EH-property unless $F$ contains a forest and the complement of a forest.

### 3.2.2 Connectivity and growing paths

We start this section with the following simple lemma.

**Lemma 3.2.6.** Let $1/2 > \varepsilon > 0$, and let $G$ be $\varepsilon$-coherent. Then $|V(G)| > 1/\varepsilon$.

**Proof.** Suppose not; let $G$ be $\varepsilon$-coherent and $n = |V(G)| \leq 1/\varepsilon$. It follows that $\varepsilon n \leq 1$. Since $G$ is $\varepsilon$-coherent, it follows that $G$ is $\varepsilon$-sparse, and therefore $|N[v]| < \varepsilon n$ for all $v \in V(G)$, and so $\varepsilon n > 1$, a contradiction. This concludes the proof.  

If a graph is $\varepsilon$-coherent, it is very “well-connected”, in the following sense.

**Lemma 3.2.7.** Let $\varepsilon > 0$, and let $G$ be an $\varepsilon$-coherent graph with $n = |V(G)|$. Then, for every $A \subseteq V(G)$ with $|A| \geq 3\varepsilon n$, there exists a connected component $X$ of $G \setminus A$ with $|V(X)| > |A| - \varepsilon n$.  

60
Proof. Let $H = G|A$, and let $C_1, \ldots, C_k$ be the connected components of $H$ such that $|V(C_1)| \geq \cdots \geq |V(C_k)|$. Let $l \in \{1, \ldots, k\}$ be the smallest number such that $|V(C_1)| + \cdots + |V(C_l)| \geq \varepsilon n$. Since $G$ has no anticomplete $(\varepsilon n, \varepsilon n)$-pair, and since $V(C_1) \cup \cdots \cup V(C_l)$ is anticomplete $V(C_{l+1}) \cup \cdots \cup V(C_k)$, it follows that $|V(C_{l+1})| + \cdots + |V(C_k)| < \varepsilon n$. This implies that $|V(C_1)| + \cdots + |V(C_l)| > |A| - \varepsilon n$.

Suppose that $l > 1$. Since $|V(C_l)| \leq |V(C_1)|$ and $|A| - \varepsilon n \geq 2\varepsilon n$, it follows that $|V(C_1)| + \cdots + |V(C_{l-1})| \geq (|V(C_1)| + \cdots + |V(C_l)|)/2 \geq \varepsilon n$, contrary to the choice of $l$. This proves that $l = 1$, and the result follows.

The proof of the following lemma follows ideas of Lemma 2.1.1 and [29]. It was first proved in [4], and it implies that Conjecture 3.1.12 holds for paths by showing that $P_k$-free graphs have the strong sparse EH-property.

Lemma 3.2.8 (Bousquet, Lagoutte, Thomassé [4]). For every $k \in \mathbb{N}$, there is an $\varepsilon > 0$ such that there is no $\varepsilon$-coherent $P_k$-free graph.

Instead of Lemma 3.2.8 we prove the following stronger statement:

Lemma 3.2.9. Let $k \in \mathbb{N}$ and $\varepsilon > 0$. Let $G$ be an $\varepsilon$-coherent graph, and let $A \subseteq V(G)$ with $|A| \geq (k + 3)\varepsilon n$ and such that $G|A$ is connected. Then, for every $v \in N(A)$, there is an induced $P_k$ in $G|A$ starting at $v$.

Proof. Let $k, \varepsilon, G, A$ be as in the statement of the lemma. Let $v \in A$. We proceed by induction on $k$. For $l \in \{1, \ldots, k\}$, we construct $v_1, \ldots, v_l$ and $A_1, \ldots, A_l$ with the following properties:

- $v_1, \ldots, v_l$ is an induced path;
- $A_l \subseteq A$ satisfies that $G|A_l$ is connected and $|A_l| > |A| - l\varepsilon n$; and
- $v_1, \ldots, v_{l-1}$ have no neighbors in $A_l$; $\{v_1, \ldots, v_l\}$ is disjoint from $A_l$, and $v_l$ has a neighbor in $A_l$. 

61
For $l = 1$, we let $v_1 = v$, and we let $A_1 = A$. Now suppose that $v_1, \ldots, v_l$ and $A_1, \ldots, A_l$ have been constructed with the properties above. If $l = k$, then $v_1, \ldots, v_l$ is the desired path and we stop.

Otherwise, it follows that $|A_l| \geq 4\varepsilon n$, and since $G$ is $\varepsilon$-sparse, it follows that $|A_l \setminus N[v_l]| \geq 3 \varepsilon n$. Moreover, we have that $|A \setminus (N[v_1] \cup \cdots \cup N[v_l])| \geq |A| - l\varepsilon n \geq 4\varepsilon n$. By Lemma 3.2.7, it follows that $G\setminus (A \setminus (N[v_1] \cup \cdots \cup N[v_l]))$ contains a connected component $A_{l+1}$ of size at least $|A| - (l + 1)\varepsilon n \geq 2\varepsilon n$. Since $|A_l \setminus N[v_l]| \geq 3 \varepsilon n$, it follows that $G\setminus (A_l \setminus N[v_l])$ contains a connected component $C$ of size at least $2\varepsilon n$. Since $A_l \setminus N[v_l] \subseteq A \setminus (N[v_1] \cup \cdots \cup N[v_l])$, it follows that $V(C) \subseteq A_{l+1}$. Since $A_l$ is connected, it follows that some vertex $v_{l+1} \in N(v_l)$ has a neighbor in $V(C) \subseteq A_{l+1}$. But then $v_{l+1}, A_{l+1}$ have the desired properties; this concludes the proof.

The following lemma proves that we can connect sets with big neighbor sets in any way we choose:

**Lemma 3.2.10.** Let $k, r \in \mathbb{N}$. Then there exists a constant $C = C(k, r) \in \mathbb{N}$ such that the following holds. Let $\varepsilon > 0$, and let $G$ be an $\varepsilon$-coherent graph. Let $X_1, \ldots, X_k, Y_1, \ldots, Y_k \subseteq V(G)$ be pairwise disjoint and pairwise anticomplete and such that $|N(X_i)|, |N(Y_i)| \geq C \varepsilon n$, and let $x_1, \ldots, x_k, y_1, \ldots, y_k$ such that $x_i \in X_i, y_i \in Y_i$ for all $i \in \{1, \ldots, k\}$. Then there exist induced paths $P_1, \ldots, P_k$ such that

- for all $i \in \{1, \ldots, k\}$, $P_i$ has one end in $X_i$ and one end in $Y_i$; and

- for all $i, j \in \{1, \ldots, k\}, i \neq j$, $V(P_i)$ is anticomplete to $X_j \cup V(P_j) \cup Y_j$.

If, in addition, it holds $X_i \subseteq N^r[x_i], Y_i \subseteq N^r[y_i]$ for all $i \in \{1, \ldots, k\}$ (i.e. $x_i$ is an $r$-center of $X_i$, and $y_i$ is an $r$-center of $Y_i$), then we can choose $P_i$ to be an $x_i$-$y_i$-path for all $i \in \{1, \ldots, k\}$.

**Proof.** We prove this by induction on $k$. For $k = 0$, the result is trivially true with $C = 1$. [Proof continues...]

62
Now suppose that $k > 0$, and that the result holds for $k - 1$ with constant $C_{k-1}$. We let $C = C_k = 2k(C_{k-1} + 2r + 5)$. Recall that by Lemma 1.3.2 if $v$ is an $r$-center for $A$, then there is a vertex $a \in A$ such that $v$ is an $r$-center for $A \setminus \{a\}$. We say that $A'$ arises from $A$ by $v$-shrinking if the following hold:

- $v \in A' \subseteq A$; and
- if $v$ is an $r$-center of $A$, then $v$ is an $r$-center of $A'$.

Then for every set $A$ and $v \in A$, either $A = \{v\}$ or there exists a vertex $a \in A$ such that $A \setminus \{a\}$ arises from $A$ by $v$-shrinking.

For $i \in \{1, \ldots, k\}$, we let $x_i \in X_i, y_i \in Y_i$ be as in the lemma. For $i \in \{1, \ldots, k - 1\}$, we let $X'_i$ be obtained from $X_i$ by $x_i$-shrinking, and let $X'_i$ be minimal such that $|N(X'_i)| \geq (C_{k-1} + 2r + 4)\varepsilon n$, and we let $Y'_i$ be obtained from $Y_i$ by $y_i$-shrinking, and let $Y'_i$ be minimal such that $|N(Y'_i)| \geq (C_{k-1} + 2r + 4)\varepsilon n$. Since either $|X'_i| = 1$ or there is a vertex $a$ in $X'_i$ such that $X'_i \setminus \{a\}$ arises from $X'_i$ by $x_i$-shrinking, it follows that $|N(X'_i)| \leq (C_{k-1}+2r+5)\varepsilon n$, and similarly, $|N(Y'_i)| \leq (C_{k-1}+2r+5)\varepsilon n$.

We now construct $P_k$. Let $Z = \bigcup_{i \in \{1, \ldots, k-1\}} (N(Y'_i) \cup N(X'_i))$. It follows that $|Z| \leq 2(k - 1)(C_{k-1} + 2r + 5)\varepsilon n$. We let $A = N(X_k) \setminus Z$ and $B = N(Y_k) \setminus Z$. It follows that $|A|, |B| \geq 2(C_{k-1} + 2r + 4)\varepsilon n$, and so $A, B$ is not an anticomplete pair. Therefore, either there exists an edge $ab \in E(G)$ with $a \in A, b \in B$, or $A \cap B \neq \emptyset$. We proceed as follows:

- If $x_k$ is not an $r$-center for $X_k$ or $y_k$ is not an $r$-center for $Y_k$, and there exists a vertex $a \in A \cap B$, then we let $x \in N(a) \cap X_k, y \in N(a) \cap Y_k$, and $P_k = x-a-y$.

- If $x_k$ is not an $r$-center for $X_k$ or $y_k$ is not an $r$-center for $Y_k$, and $A \cap B = \emptyset$, then we let $ab \in E(G)$ with $a \in A, b \in E$, and we let $x \in N(a) \cap X_k, y \in N(b) \cap Y_k$, and $P_k = x-a-b-y$.
• If \( x_k \) is an \( r \)-center for \( X_k \) and \( y_k \) is an \( r \)-center for \( Y_k \), then

\[
d_{G|(X_k \cup Y_k \cup A \cup B)}(x_k, y_k) \leq 2r + 3.
\]

We let \( P_k \) be a shortest \( x_k-y_k \)-path in \( G|(X_k \cup Y_k \cup A \cup B) \).

By definition, and since \( V(P_k) \) is disjoint from \( Z \), it follows that \( X'_j, Y'_j \) are anti-
complete to \( V(P_k) \) for all \( j \in \{1, \ldots, k-1\} \). We let \( G' = G \setminus (N[V(P_k)]) \). It follows
that

\[
|N(X'_i) \cap V(G')| \geq |N(X'_i)| - |V(P_k)|\varepsilon n \geq (C_{k-1} + 2r + 4)\varepsilon n - (2r + 4)\varepsilon n \geq C_{k-1}\varepsilon n.
\]

It follows that \( X'_1, \ldots, X'_k, Y'_1, \ldots, Y'_k \subseteq V(G') \) are pairwise disjoint and pairwise
anticonplete and satisfy that \( |N_G'(X'_i)|, |N_G'(Y'_i)| \geq C_{k-1}\varepsilon n \), and \( x_1, \ldots, x_k, y_1, \ldots, y_k \)
satisfy that \( x_i \in X_i, y_i \in Y_i \) for all \( i \in \{1, \ldots, k\} \), and if \( x_i \) is an \( r \)-center for \( X_i \) for all
\( i \in \{1, \ldots, k\} \) and \( y_i \) is an \( r \)-center for \( Y_i \) for all \( i \in \{1, \ldots, k\} \), then \( x_i \) is an \( r \)-center
for \( X'_i \) and \( y_i \) is an \( r \)-center for \( Y'_i \) for all \( i \in \{1, \ldots, k\} \). By induction, it follows that
we can find \( P_1, \ldots, P_{k-1} \) as desired. By adding \( P_k \), which is anticonplete to \( V(G') \)
and hence to \( V(P_1), \ldots, V(P_{k-1}) \), the result of the lemma follows. \( \Box \)

### 3.2.3 Caterpillars

A tree \( T \) is a caterpillar if there is a path \( P \) in \( T \) such that every vertex in \( V(T) \setminus V(P) \)
is a leaf of \( T \). In this section, we prove Conjecture [3.2.3] for caterpillars.

Let \( T \) be a tree and \( v \in V(T) \). We denote by \( T^v \) the directed graph arising from
\( T \) by directing an edge \( uw \) as \((u, w)\) if and only if \( w \) is a vertex of the unique \( u-v \)-path
in \( T \). A \((T, v)\)-cover structure is a collection of sets \( \{A_t\}_{t \in V(T)} \) such that

• for every edge \((u, w) \in E(T^v)\), \( A_u \) covers \( A_w \); and

• for all \( u, w \in V(T) \) with \( u \neq w \), the set \( A_u \) is anticonplete to \( A_w \) if \( uw \not\in E(T) \).
The following lemma shows how a \((T,v)\)-cover structure in \(G\) can be used to find an induced copy of \(T\) in \(G\).

**Lemma 3.2.11.** Let \(G\) be a graph that contains a \((T,v)\)-cover structure \(\{A_t\}_{t \in V(T)}\) with \(A_v \neq \emptyset\). Then \(G\) contains \(T\) as an induced subgraph, and moreover, \(G\) contains a copy \(T'\) of \(T\) such that for all \(t \in V(T)\), the vertex of \(T'\) corresponding to \(t\) is contained in \(A_t\).

**Proof.** We define a function \(f : V(T) \rightarrow V(G)\) as follows. We let \(f(v) \in A_v\) be chosen arbitrarily. Then, for every edge \((u,v) \in E(T^v)\), \(A_u\) covers \(A_v\), and thus \(v\) has a neighbor in \(A_u\). We let \(f(u) \in A_u \cap N(f(v))\). We continue this way; if there is an edge \((u,w) \in E(T^w)\) such that \(f(w)\) is defined and \(f(u)\) is not, then we know that \(f(w) \in A_w\), and \(A_u\) covers \(A_w\); thus we let \(f(u) \in A_u \cap N(f(w))\). Since every vertex \(u \in V(T^v)\) has out-degree one in \(T^v\) and starts a directed path that ends in \(v\), it follows that for every vertex \(u\), \(f(u)\) is defined exactly once.

For every edge \((u,w) \in E(T^w)\), since \(w\) is the only out-neighbor of \(u\), it follows that \(f(u)\) was defined to be adjacent to \(f(w)\). It follows that \(f(u)f(w) \in E(G)\). Suppose that there exists \(u,w\) with \(f(u)f(w) \in E(G)\), but \(uw \notin E(T)\). Then \(A_u\) is anticomplete to \(A_w\) by the definition of a \((T,v)\)-cover structure, but \(f(u) \in A_u\) and \(f(w) \in A_w\), a contradiction. It follows that \(T' = G|\{f(u) : u \in V(T)\}\) is isomorphic to \(T\) and has the desired properties. This concludes the proof. \(\square\)

We need the following definitions. Let \(T\) be a caterpillar, and let \(P\) be a path in \(T\) containing all vertices of degree more than one. Let \(v\) be an end of \(P\). Then \(v\) is called a **head** of \(T\), and \(P\) is called a **spine** for \(v\) in \(T\). Let \(T\) be a caterpillar, let \(v\) be a head of \(T\) and let \(P\) be a spine for \(v\) in \(T\). Let \(P = v_1 \ldots v_k\), where \(v = v_k\). Then a \((T,v)\)-caterpillar ordering is an ordering \(\prec\) of \(V(T)\) in which the following hold:

- for all \(i \in \{1, \ldots, k\}\), \(v_i \prec w\) for all \(w \in N_T(v_i) \setminus V(P)\);
- for all \(i, j \in \{1, \ldots, k\}\), \(v_i \prec v_j\) if and only if \(i < j\); and
• for all \( i \in \{1, \ldots, k-1\} \), \( w \prec v_{i+1} \) for all \( w \in N_T(v_i) \setminus V(P) \).

For a \((T, v)\)-caterpillar ordering \( \prec \), we write \( V(T) = \{w_1, \ldots, w_t\} \) with \( w_1 \prec \cdots \prec w_t \).

We say that \((T', v')\) is a \((T, v, \prec)\)-initial segment if \( T' = T|\{w_1, \ldots, w_i\} \) for some \( i \in \{1, \ldots, t\} \), and if \( v' = v_j \) where \( j = \{\max i : v_i \in V(T')\} \) is the last element of the spine of \( T \) in \( V(T') \).

For \( c > 0 \) and \( k \in \mathbb{N} \), a \((k, c)\)-box array in a graph \( G \) is a collection \( B_1, \ldots, B_k \) of disjoint subsets of \( V(G) \) with \( |B_i| \geq c|V(G)| \) for all \( i \in \{1, \ldots, k\} \). A \((T, v)\)-cover structure \( \{A_t\}_{t \in V(T)} \) is \( \mathcal{B}\)-rainbow for a \((k, c)\)-box array \( \mathcal{B} = B_1, \ldots, B_k \) if for every \( t \in V(T) \), there exists a \( j \in \{1, \ldots, k\} \) such that \( A_t \subseteq B_j \), and for every \( j \in \{1, \ldots, k\} \), there is at most one \( t \in V(T) \) such that \( A_t \cap B_j \neq \emptyset \).

For \( c, k \geq 0 \), a caterpillar \( T \), a head \( v \) of \( T \) and a \((T, v)\)-caterpillar ordering \( \prec \), a \((k, c, T, v, \prec)\)-caterpillar array is a \((k, c)\)-box array \( B_1, \ldots, B_k \) and a list \( \{A^1_{t_i}\}_{t \in V(T)} \ldots \{A^k_{t_i}\}_{t \in V(T)} \) of initial caterpillars such that

- for every \( i \in \{1, \ldots, k\} \), \((T_i, t_i)\) is a \((T, v, \prec)\)-initial segment and \( \{A^i_{t_i}\}_{t \in V(T_i)} \) is a \((T_i, t_i)\)-cover structure;

- \( A^i_{t_i} = B_i \); and

- for \( i, j \in \{1, \ldots, k\} \) with \( i \neq j \), and for all \( t \in V(T_i), t' \in V(T_j) \), such that either \( t \neq t_i \) or \( t' \neq t_j \), we have that \( A^i_{t_i} \) is anticomplete to \( A^j_{t_j} \).

The following lemma is a “black box”: we will apply this lemma with different procedures \( \mathcal{P} \) to generate sets \( A \) as in Lemma 3.2.12 with additional properties, which will allow us to prove strengthenings of Lemma 3.2.14

**Lemma 3.2.12.** Let \( \varepsilon > 0 \), and let \( G \) be an \( \varepsilon \)-coherent graph. Let \( n = |V(G)| \). Let \( c \geq 2\varepsilon \), and let \( k \in \mathbb{N} \) with \( k \geq 2 \). Let \( \mathcal{B} = (B_1, \ldots, B_k) \), where \( B_1, \ldots, B_k \) is a \((k, c)\)-box array in \( G \), and let \( i \in \{1, \ldots, k\} \).

Then there exists a \((k-1, c/2 - \varepsilon)\)-box array \( \mathcal{B}' = (B_1', \ldots, B_{i-1}', B_i', \ldots, B_k') \) and a set \( A \subseteq B_i \), and a \( j \in \{1, \ldots, k\} \setminus \{i\} \) such that
• A covers \( B'_j \);

• for all \( l \in \{1, \ldots, k\} \setminus \{i, j\} \), \( A \) is anticomplete to \( B'_l \); and

• for all \( l \in \{1, \ldots, k\} \setminus \{i\} \), \( B'_l \subseteq B_i \).

Additionally, if \( \mathcal{P} \) is a procedure that, given \( B_i \) and \( B \), outputs a set \( \{a_1, \ldots, a_t\} = B'_i \subseteq B_i \) such that

\[
\left| (B_1 \cup \cdots \cup B_k) \setminus (B_i \cup N(B'_i)) \right| < \frac{\left| (B_1 \cup \cdots \cup B_k) \setminus B_i \right|}{2},
\]

then we may choose \( A \) such that \( A = \{a_1, \ldots, a_s\} \) for some \( s \in \{1, \ldots, t\} \).

In particular, if \( \mathcal{P} \) is a procedure that outputs a set \( B'_i \subseteq B_i \), and a labeling \( \{a_1, \ldots, a_t\} = B'_i \) such that \( |V(G) \setminus N(B'_i)| < \varepsilon n \) (e.g. by choosing \( B'_i \) with \( |B'_i| \geq \varepsilon n \)), then we may choose \( A \) such that \( A = \{a_1, \ldots, a_s\} \) for some \( s \in \{1, \ldots, t\} \).

**Proof of Lemma 3.2.12** Let \( G, \varepsilon, n, \mathcal{B}, k, c \) and \( i \) be as in the statement of the lemma. If we are given a procedure \( \mathcal{P} \), we let \( B'_i \) be obtained by applying \( \mathcal{P} \) to \( B_i \) and \( B \). Otherwise, we let \( B'_i = B_i \), and note that since \( |B_i| \geq cn \geq \varepsilon n \), and since \( G \) is \( \varepsilon \)-coherent, it follows that \( B'_i \) satisfies that

\[
\left| (B_1 \cup \cdots \cup B_k) \setminus (B_i \cup N(B'_i)) \right| < \varepsilon n \leq \frac{\left| (B_1 \cup \cdots \cup B_k) \setminus B_i \right|}{2}.
\]

In both cases (whether \( \mathcal{P} \) was given or not), it follows that \( |(B_1 \cup \cdots \cup B_k) \setminus (B_i \cup N(B'_i))| < |(B_1 \cup \cdots \cup B_k) \setminus B_i|/2 \). We let \( B'_i = \{a_1, \ldots, a_t\} \), where this labeling is the output of \( \mathcal{P} \) if a procedure \( \mathcal{P} \) was given, and the labeling is chosen arbitrarily otherwise.

We let \( A \subseteq B'_i \) be such that \( A = \{a_1, \ldots, a_s\} \), where we let \( s \in \{1, \ldots, t\} \) be maximal subject to the condition that \( |N(A) \cap B_j| \leq cn/2 \) for all \( j \in \{1, \ldots, k\} \setminus \{i\} \).
Since

$$(k - 1)cn/2 \leq \left|\frac{(B_1 \cup \cdots \cup B_k) \setminus B_i}{2}\right| < \left|\left((B_1 \cup \cdots \cup B_k) \setminus B_i\right) \cap N(B'_i)\right|,$$

it follows that $A \neq B'_i$. Let $a = a_{s+1} \in B'_i \setminus A$. By the choice of $s$, it follows that there exists a $j \in \{1, \ldots, k\} \setminus \{i\}$ such that $|N(A \cup \{a\}) \cap B_j| > cn/2$. Since $|N(a)| \leq \varepsilon n$, it follows that $|N(A) \cap B_j| \geq cn/2 - \varepsilon n$. Now we let

- $B'_j = N(A) \cap B_j$; and
- $B'_l = B_l \setminus N(A)$ for all $l \in \{1, \ldots, k\} \setminus \{i, j\}$.

It follows that $|B'_l| \geq |B_l| - |N(A) \cap B_l| \geq cn - cn/2 \geq cn/2$ for all $l \in \{1, \ldots, k\} \setminus \{i, j\}$. Consequently, $\mathcal{B}' = (B'_1, \ldots, B'_{i-1}, B'_{i+1}, \ldots, B'_k)$ is a $(k - 1, c/2 - \varepsilon)$-box array, and $A$ covers $B'_j$ and is anticomplete to $B'_l$ for all $l \in \{1, \ldots, k\} \setminus \{i, j\}$, as claimed. This concludes the proof.

One particularly useful application of Lemma 3.2.12 is the following.

**Lemma 3.2.13.** Let $\varepsilon > 0$, and let $G$ be an $\varepsilon$-coherent graph. Let $n = |V(G)|$. Let $k \in \mathbb{N}$ be such that $k \geq 2$, and let $c \geq 3\varepsilon$. Let $\mathcal{B} = (B_1, \ldots, B_k)$, where $B_1, \ldots, B_k$ is a $(k, c)$-box array in $G$, and let $i \in \{1, \ldots, k\}$. Then there exists a $(k - 1, c/2 - \varepsilon)$-box array $\mathcal{B}' = (B'_1, \ldots, B'_{i-1}, B'_{i+1}, \ldots, B'_k)$, a set $A \subseteq B_i$ such that $G|A$ is connected, and a $j \in \{1, \ldots, k\} \setminus \{i\}$ such that

- $A$ covers $B'_j$;
- for all $l \in \{1, \ldots, k\} \setminus \{i, j\}$, $A$ is anticomplete to $B'_l$; and
- for all $l \in \{1, \ldots, k\} \setminus \{i\}$, $B'_l \subseteq B_l$.

**Proof.** It follows from Lemma 3.2.7 that $B_i$ contains a subset $B'_i$ such that $G|B'_i$ is connected and $|B'_i| \geq |B_i| - \varepsilon n$. By Lemma 1.3.3 we can choose a labeling
\{a_1, \ldots, a_t\} = B'_t \text{ such that } G|\{a_1, \ldots, a_s\} \text{ is connected for all } s \in \{1, \ldots, t\}. \text{ Now the result follows from Lemma 3.2.12.} \qed

The following lemma is the main technical result of this section.

**Lemma 3.2.14.** • Let \( T \) be a caterpillar with head \( v \) and let \( C, \varepsilon > 0 \). Then there exist \( c, K, C' > 0 \) with \( c \geq C \varepsilon \) such that for every \( (K, C' \varepsilon) \)-box array \( \mathcal{B} \), and for every \( \varepsilon \)-coherent graph \( G \), \( G \) contains a \( \mathcal{B} \)-rainbow \( (T,v) \)-cover structure \( \{A_t\}_{t \in V(T)} \) with \( |A_v| \geq cn \), where \( n = |V(G)| \).

• Let \( T \) be a caterpillar with head \( v \) and let \( C > 0 \). Then there exist \( \varepsilon, c > 0 \) with \( c \geq C \varepsilon \) such that for every \( \varepsilon \)-coherent graph \( G \), \( G \) contains a \( (T,v) \)-cover structure \( \{A_t\}_{t \in V(T)} \) with \( |A_v| \geq cn \), where \( n = |V(G)| \).

Moreover, for both statements, for every \( t \in V(T) \setminus \{v\} \), the set \( A_t \) is chosen according to Lemma 3.2.12 applied to a \( (k',c') \)-box array with \( c' \geq c \) and \( k' \geq 2 \), and \( A_t \) satisfies that \( |N_G(A_t)| \geq cn \) and \( |A_t| \leq \varepsilon n \).

**Proof.** We will prove both statements simultaneously. We first describe a procedure for transforming caterpillar arrays. Throughout this proof, we let \( T \) be a caterpillar and we let \( v \) be a head for \( T \). We fix a spine \( P \) for \( T \) and \( v \), and a \( (T,v) \)-caterpillar ordering \( \prec \). We let \( V(T) = \{w_1, \ldots, w_t\} \) such that \( w_1 \prec \cdots \prec w_t \).

Now let \( G \) be a graph; let \( n = |V(G)| \). Let \( k \in \mathbb{N} \) and \( c > 0 \). For a \( (k,c,T,v,\prec) \)-caterpillar array \( \mathcal{A} \) with initial caterpillars \( \{A^1_t\}_{t \in V(T_1)}, \ldots, \{A^k_t\}_{t \in V(T_k)} \), we define \( \varphi(\mathcal{A}) = \sum_{i=1}^{k} 2^{|V(T_i)|} \).

We need the following definitions. Let \( \mathcal{A} \) be a \( (k,c,T,v,\prec) \)-caterpillar array with initial caterpillars \( \{A^1_t\}_{t \in V(T_1)}, \ldots, \{A^k_t\}_{t \in V(T_k)} \). Then \( \mathcal{A} \) is **incomplete** if \( |V(T_i)| < |V(T)| \) for all \( i \in \{1, \ldots, k\} \). We call an initial caterpillar \( \{A^i_t\}_{t \in V(T_i)} \) **saturated** if
\[ V(T_i) = \{w_1, \ldots, w_j\} \] and either \(|V(T)| = |V(T_i)|\) or \(w_{j+1} \in V(P)\), where \(P\) is the spine we chose for \(T\).

Let \(\varepsilon > 0\), and let \(G\) be \(\varepsilon\)-coherent. If \(G\) contains an incomplete \((k, c, T, v, \prec)\)-caterpillar array \(A\) with \(k > 1\), then \(G\) contains a \((k - 1, c/2 - \varepsilon, T, v, \prec)\)-caterpillar array \(A'\) with \(\varphi(A') \geq \varphi(A)\).

To prove this, we let \(k \in \mathbb{N}\) with \(k > 1\), and we let \(c > 0\). Let \(A\) be an incomplete \((k, c, T, v, \prec)\)-caterpillar array consisting of a \((k, c)\)-box array \(B_1, \ldots, B_k\) and initial caterpillars \(\{A^1_1\}_{t \in V(T_1)}, \ldots, \{A^k_k\}_{t \in V(T_k)}\). By symmetry, we may assume that \(|V(T_1)| \geq \cdots \geq |V(T_k)|\). We now choose a number \(i \in \{1, \ldots, k\}\), as follows:

- If no initial caterpillar in \(\{A^1_1\}_{t \in V(T_1)}, \ldots, \{A^k_k\}_{t \in V(T_k)}\) is saturated, we let \(i = k\).
- Otherwise, we let \(i\) be the minimum \(j \in \{1, \ldots, k\}\) such that \(\{A^j_j\}_{t \in V(T_j)}\) is saturated.

Now we apply Lemma 3.2.12 to \(B_1, \ldots, B_k\) and with index \(i\). We obtain a \((k - 1, c/2 - \varepsilon)\)-box array \(B'_1, \ldots, B'_{i-1}, B'_{i+1}, \ldots, B'_k\), a set \(A \subseteq B_i\), and a \(j \in \{1, \ldots, k\} \setminus \{i\}\) such that

- \(A\) covers \(B'_j\);
- for all \(l \in \{1, \ldots, k\} \setminus \{i, j\}\), \(A\) is anticomplete to \(B'_l\); and
- for all \(l \in \{1, \ldots, k\} \setminus \{i\}\), \(B'_l \subseteq B_l\).

Since \(G\) is \(\varepsilon\)-coherent, it follows that \(|A| \leq \varepsilon n\). We consider two cases:

Suppose first that \(i > j\). By the choice of \(i\), it follows that \(\{A^j_j\}_{t \in V(T_j)}\) is not saturated. Let \(V(T_j) = \{w_1, \ldots, w_s\}\). Since \(A\) is incomplete, it follows that \(|V(T_j)| < |V(T)|\). Therefore, \(w_{s+1}\) exists and is not in the spine \(P\) of \(T\). Let \(t \in V(T)\) be such that \(A^j_j = B_j\). By the definition of a caterpillar ordering, and since \((T_j, t)\) is a \((T, v, \prec)\)-initial segment, it follows that \(w_{s+1} \in N_T(t)\). We let \(T'_j = T|\{w_1, \ldots, w_{s+1}\}\), and we
let $A^i_t = B'_t$, $A^i_{w_{s+1}} = A$, and $A^i_{t'} = A_{t'}$ for all $t' \in V(T_j) \setminus \{t\}$. Since $A \subseteq B_i$, it follows that $A$ is anticomplete to $A^i_t$ for all $t' \in V(T_j) \setminus \{t\}$, and by Lemma 3.2.12 it follows that $A^i_{w_{s+1}}$ covers $A^i_t$. This implies that $\{A^i_{t'}\}_{t' \in V(T_j)}$ is a $(T'_j, t)$-cover structure.

For $l \in \{1, \ldots, k\} \setminus \{i, j\}$, we define the following:

- we let $t \in V(T_l)$ such that $(T_l, t)$ is a $(T, v, \prec)$-initial segment;
- we let $A^i_t = B'_t$;
- we let $A^i_{t'} = A^i_{t'}$ for all $t' \in V(T_l) \setminus \{t\}$; and
- we let $T'_l = T_l$.

The $(k-1, c/2-\varepsilon)$-box array $B'_1, \ldots, B'_{i-1}, B'_{i+1}, \ldots, B'_k$ together with the initial caterpillars $\{A^i_1\}_{t \in V(T_1)}, \ldots, \{A^i_{t-1}\}_{t \in V(T_{l-1})}$, $\{A^i_{t+1}\}_{t \in V(T_{l+1})}$ is a $(k-1, c/2-\varepsilon, T, v, \prec)$-caterpillar array $A'$. Moreover,

$$\varphi(A') = \varphi(A) - 2^{|V(T_i)|} - 2^{|V(T_j)|} + 2^{|V(T_j)|} \geq \varphi(A)$$

since $|V(T_j)| = |V(T_j)| + 1$ and $|V(T_i)| \leq |V(T_j)|$. This proves that (3.1) holds in the case when $i > j$.

Now suppose that $i < j$. By the choice of $i$, it follows that $\{A^i_t\}_{t \in V(T_i)}$ is saturated. Let $V(T_i) = \{w_1, \ldots, w_s\}$. Since $A$ is incomplete, it follows that $|V(T_i)| < |V(T)|$. Therefore, $w_{s+1}$ is a vertex of the spine $P$ of $T$. By the definition of a caterpillar ordering, and since $(T_i, t)$ is a $(T, v, \prec)$-initial segment, it follows that $w_{s+1} \in N_T(t)$, and that $(T \setminus \{w_1, \ldots, w_{s+1}\} \cup w_{s+1})$ is a $(T, v, \prec)$-initial segment. We let $T'_j = T \setminus \{w_1, \ldots, w_{s+1}\}$, and we let $A^i_{w_{s+1}} = B'_j$, $A^i_j = A$ and $A^i_{t'} = A^i_{t'}$ for all $t' \in \{w_1, \ldots, w_s\} \setminus \{t\}$. Since $A^i_t = B_i$, it follows that $B_j$ is anticomplete to $A^i_{t'}$ for all $t' \in V(T_j) \setminus \{t\}$. Moreover, $A^i_j$ covers $B_i$ and thus $A = A^i_{w_{s+1}}$. This implies that $\{A^i_{t'}\}_{t' \in V(T'_j)}$ is a $(T'_j, w_{s+1})$-cover structure.

For $l \in \{1, \ldots, k\} \setminus \{i, j\}$, we define the following:
• we let \( t \in V(T_i) \) such that \((T_i, t)\) is a \((T, v, \prec)\)-initial segment;
• we let \( A_i^t = B_i' \);
• we \( A_i^t = A_i' \) for all \( t' \in V(T_i) \setminus \{t\} \); and
• we let \( T_i' = T_i \).

It follows that the \((k - 1, c/2 - \varepsilon)\)-box array \( B_1', \ldots, B_{i-1}', B_{i+1}', \ldots, B_k' \) together with initial caterpillars \( \{A_i^t\}_{t \in V(T_i')} \), \( \ldots \), \( \{A_{i-1}^{t-1}\}_{t \in V(T_{i-1}')} \), \( \{A_i^{t+1}\}_{t \in V(T_{i+1}')} \) is a \((k - 1, c/2 - \varepsilon, T, v, \prec)\)-caterpillar array \( \mathcal{A}' \). Moreover,

\[
\varphi(\mathcal{A}') = \varphi(\mathcal{A}) - 2^{|V(T_i')|} - 2^{|V(T_j')|} + 2^{|V(T_j')|} \geq \varphi(\mathcal{A}),
\]

since \( |V(T_j')| = |V(T_i)| + 1 \) and \( |V(T_i)| \geq |V(T_j)| \). This proves that (3.1) when \( i < j \).

This concludes the proof of (3.1).

We now return to the proof of the lemma. For proving the first statement, we let \( K = 2^{|V(T)|} \), and \( C' = C \cdot 4^{K+1} \). For proving the second statement, we let \( K = 2^{|V(T)|} \), and let \( \varepsilon \leq (K + 1)^{-2}4^{-K-1} \cdot (1/C) \). It follows that \( \varepsilon \leq 1/(K + 1)^2 \), and so \( n \geq 1/(K + 1)^2 \) by Lemma 3.2.6. This implies that \( n/K - 1 \geq Kn/(K(K + 1)) + n/(K + 1)) - 1 \geq n/(K + 1) \). We define a \((K, 1/(K + 1))\)-box array \( B_1, \ldots, B_k \) by partitioning \( V(G) \) into disjoint sets of size at least \( |n/K| \geq n/K - 1 \) arbitrarily.

Now, for all \( i \in \{1, \ldots, K\} \), we let \( T_i = T \setminus \{w_i\} \), and we let \( A_{w_i}^i = B_i \). Then \( B_1, \ldots, B_K \) and \( \{A_{w_i}^i\}_{i \in V(T_1)}, \ldots, \{A_{w_K}^K\}_{i \in V(T_K)} \) is a \((K, 1/(K + 1), T, v, \prec)\)-caterpillar array \( \mathcal{A} \), and \( \varphi(\mathcal{A}) = K \cdot 2 = 2K \).

We let \( \mathcal{A}_0 = \mathcal{A} \). We recursively apply (3.1) to \( \mathcal{A}_i \) and call the resulting \((K - i, c_i, T, v, \prec)\)-caterpillar array \( \mathcal{A}_{i+1} \); if \( \mathcal{A}_{i+1} \) is not incomplete or \( i = K - 1 \), we stop. This procedure terminates after at most \( K \) steps, since it reduces the number of boxes in the box array. We let \( \mathcal{A}' \) denote the \((k', c', T, v, \prec)\)-caterpillar array we obtain at termination; by construction, we have that \( \varphi(\mathcal{A}') \geq \varphi(\mathcal{A}) = 2K \).
We prove this by induction; it holds for $i = 1$. For $i > 1$, we have that
\[
c_i + 1 \geq c_i / 2 - \varepsilon \geq C4^{K+2-i}/2 - \varepsilon \geq C4^{K+2-i-1}\varepsilon.
\]
This implies that $c' \geq C\varepsilon$, as claimed.

If we are proving the second statement, we have that $c_1 \geq 1/(2(K + 1)) - \varepsilon$, and since $1/(2(K + 1)) \geq 2\varepsilon$, it follows that $c_1 \geq 1/(4(K + 1))$. For $i > 1$, by induction, we may assume that $c_i \geq 1/(K + 1) \cdot 1/i$; it follows that $c_i \geq 2\varepsilon$ and $c_{i+1} \geq c_i / 2 - \varepsilon \geq c_i / 4 = 1/(K + 1) \cdot 1/i$. This implies that $c' \geq 1/(K + 1) \cdot 4^{-1} \geq C\varepsilon$, and thus our claim follows.

Let $\{A^i_t\}_{t \in V(T_i)}$ be the initial caterpillars of $\mathcal{A}'$. If $\mathcal{A}'$ is not incomplete, it follows that there exists an $i \in \{1, \ldots, k'\}$ such that $|V(T_i)| = |V(T)|$; but then $\{A^i_t\}_{t \in V(T_i)}$ is a $(T, v)$-cover structure with $A^i_v = B_i$ and thus $|A^i_v| \geq c'n \geq C\varepsilon n$, as claimed. Therefore, we may assume that $k' = 1$, and the initial caterpillars of $\mathcal{A}'$ consist of only $\{A^1_t\}_{t \in V(T_1)}$. We have that $\varphi(\mathcal{A}') \geq \varphi(\mathcal{A}) = 2K$, but $\varphi(\mathcal{A'}) = 2^{|V(T_1)|}$, and so $2^{|V(T_1)|} \geq 2K = 2 \cdot 2^{|V(T)|}$. This is a contradiction, since $|V(T)| \geq |V(T_i)|$ by the definition of a $(k', c', T, v, \prec)$-caterpillar array. This proves that this case does not happen, and the result follows.

We deduce the main result of this section.

**Theorem 3.2.15.** Let $T$ be a caterpillar. Then there exists an $\varepsilon > 0$ such that every $T$-free $\varepsilon$-sparse graph $G$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$.

**Proof.** Let $v$ be a head for $T$. By Lemma 3.2.14 there exists an $\varepsilon > 0$ such that every $\varepsilon$-sparse graph $G$ with no anticomplete $(\varepsilon n, \varepsilon n)$-pair contains a $(T, v)$-cover structure $\{A_t\}_{t \in V(T)}$ with $|A_v| \geq \varepsilon n > 0$, where $n = |V(G)|$. By Lemma 3.2.11 it follows that if $G$ contains a $(T, v)$-cover structure $\{A_t\}_{t \in V(T)}$ with $A_v \neq \emptyset$, then $G$ contains $T$ as
an induced subgraph. Thus it follows that every $T$-free $\varepsilon$-sparse graph $G$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$. This implies the statement of the theorem.

3.2.4 Caterpillar subdivisions

The proof of the following lemma uses the fact that while for any given $k$ and $\varepsilon > 0$, there is a $\varepsilon$-coherent $C_k$-free graph by Lemma 3.2.5, the same conclusion does not hold if we consider $\{C_l : l \geq k\}$-free graphs [3]:

Lemma 3.2.16 (Bonamy, Bousquet, Thomassé [3]). For all $k$, the class $\mathcal{C}$ of graph with no holes of length at least $k$ and no antiholes of length at least $k$ has the strong EH-property.

The main result of this section is the following generalization of Lemma 3.2.16:

Theorem 3.2.17. For every graph $H$, there exists an $\varepsilon > 0$ such that every $\varepsilon$-coherent graph $H$ contains an induced subgraph $H'$ isomorphic to a subdivision of $H$.

We will prove the following stronger statement:

Theorem 3.2.18. For every graph $H$, and for every (not necessarily induced) path $P$ in $H$, there exists an $\varepsilon > 0$ such that for every graph $G$, one of the following holds:

- $G$ contains an induced subgraph $H'$ such that $H'$ is isomorphic to a subdivision of $H$ in which precisely those edges of $H$ corresponding to the edges in $P$ are not subdivided;

- $G$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$; or

- there is a vertex $v \in V(G)$ with $|N[v]| \geq \varepsilon |V(G)|$.

We call a subdivision $H'$ of $H$ as in the first bullet a $P$-subdivision of $H$. A (not necessarily induced) path $P$ in a graph $H$ is a Hamilton path if $V(H) = V(P)$. Note
that to prove the above, we may assume that $P$ is a Hamilton path: If $V(H) = \{v_1, \ldots, v_k\}$ and $V(P) = \{v_1, \ldots, v_l\}$, for $l < k$, we add $k - l$ vertices $w_{l+1}, \ldots, w_k$ and edges $v_{i-1}w_i$ and $w_iv_i$ for $i \in \{l+1, \ldots, k\}$ to $H$ to obtain a new graph $H'$, and let $P' = Pv_l-w_{l+1}-v_{l+1}-\ldots-w_k-v_k$. Then every $P'$-subdivision of $H'$ contains a $P$-subdivision of $H$ as an induced subgraph.

The proof of Theorem 3.2.18 is divided into two parts, proved separately in Lemma 3.2.19 and Lemma 3.2.22. We say that a graph $G$ is $(\varepsilon, r)$-coherent if

- $|N^r[v]| \leq \varepsilon|V(G)|$ for all $v \in V(G)$; and
- $G$ has no anticomplete $(\varepsilon|V(G)|, \varepsilon|V(G)|)$-pair.

It follows that every $\varepsilon$-coherent graph is $(\varepsilon, 1)$-coherent and vice versa.

**Lemma 3.2.19.** Let $H$ be a graph, and let $P$ be a Hamilton path in $H$. Let $r \geq 1$, and let $c > 0$. Then there exists an $\varepsilon > 0$ such that if there is a graph $G$ that does not contain a $P$-subdivision of $H$ and $G$ is $(\varepsilon, r)$-coherent, then there is a graph $G'$ that does not contain a $P$-subdivision of $H$ and $G'$ is $(c, r+1)$-coherent.

**Proof.** Let $H$, $p$, $r$ and $c$ be as in the statement of the lemma. Let $C = C(\|E(H) \setminus E(P)\|, 2r + 2)$ as in Lemma 3.2.10 and let $C' = C + 3|V(H)| + 6|E(H)|$. Let $T$ be a caterpillar with spine $P$, head $v$ (where $v$ is an end of $P$), and in which every vertex $p \in V(P)$ has degree $d_H(p)$. We define a map $f : V(T) \setminus V(P) \to E(H) \setminus E(P)$ such that for every edge $e \in E(H) \setminus E(P)$, $f^{-1}(e) = \{x, y\}$, where $N_T(x) \cap V(P) = \{u\}$, $N_T(y) \cap V(P) = \{v\}$, and $e = uv$.

By Lemma 3.2.14, it follows that there exist $d, e > 0$ such that $d \geq C' \cdot e$ and every $e$-coherent graph $G'$ contains a $(T, v)$-cover structure $\{A_t\}_{t \in V(T)}$ with $|A_t| \geq dn$, where $n = |V(G)|$. Moreover, every $A_t$ with $t \in V(T) \setminus \{v\}$ is chosen according to Lemma 3.2.12 applied to a $(k', c')$-box array with $c' \geq d$, and $|N_G(A_t)| \geq dn$. 

75
We may assume that \( e \leq c \). Let \( \varepsilon = c \cdot e \). Let \( G \) be a \((\varepsilon, r)\)-coherent graph that does not contain a \( P \)-subdivision of \( H \), and let \( n = |V(G)| \). Suppose that the statement of the lemma does not hold for \( G \) and \( \varepsilon \).

Let \( X \subseteq V(G) \) with \( |X| \geq en \). Then there is a vertex \( v \in V(G) \) such that \( |N_{G[X]}^r[v]| \geq c|X| \geq \varepsilon|V(G)| \).

(3.2)

Suppose not. Since \( |X| \geq (\varepsilon/c)n \), it follows that \( G \cdot X \) is \((\varepsilon/c)\)-coherent. Since \( e \leq c \), it follows that \( G \cdot X \) is \( c \)-coherent. Therefore, if \( |N_{G[X]}^{r+1}[v]| < c|X| \) for all \( v \in X \), it follows that \( G \cdot X \) is \((c, r+1)\)-coherent, contrary to our assumption that the statement of the lemma does not hold for \( G \) and \( \varepsilon \). This proves (3.2).

When constructing \( \{A_t\}_{t \in V(T)} \) using Lemma 3.2.14, we may assume that for every \( t \in V(T) \setminus \{v\} \), \( A_t \) is chosen by applying Lemma 3.2.12 to a set \( B_t \) with \( |B_t| \geq dn \geq en = (\varepsilon/c)n \). We define a procedure \( \mathcal{P} \) for Lemma 3.2.12. It follows from (3.2) that \( B_t \) contains a vertex \( w \) with \( |N_{G[B_t]}^r[w]| \geq \varepsilon n \). Let \( B_t' = N_{G[B_t]}^r[w] \) and let \( \{a_1, \ldots, a_s\} \) be a labeling of \( B_t' \) such that for all \( i \in \{0, \ldots, r\} \) there is a \( j \in \{1, \ldots, s\} \) such that \( \{a_1, \ldots, a_j\} = N_{G[B_t]}^i[w] \). This implies that we may assume that every \( A_t \) with \( t \in V(T) \setminus \{v\} \) contains a vertex \( v_t \in A_t \) such that \( A_t = N_{G[A_t]}^{r+1}[v_t] \).

Now, by Lemma 3.2.14, we let \( \{A_t\}_{t \in V(T)} \) be a \((T, v)\)-cover structure such that

- For every \( t \in V(T) \setminus \{v\} \), there exists a vertex \( v_t \in A_t \) such that \( A_t = N_{G[A_t]}^{r+1}[v_t] \), \( |N_G(A_t)| \geq dn \) and \( |A_t| \leq \varepsilon n \); and
- \( |A_v| \geq dn \).

We let \( P = v-p_1-\ldots-p_k \). We let \( x(v) \in A_v \), set \( p_0 = v \), and for \( i \in \{1, \ldots, k\} \), we let \( x(p_i) \in A_{p_i} \cap N_G(x(p_{i-1})) \). For all \( i \in \{0, \ldots, k\} \) and for all \( w \in N_T(p_i) \setminus V(P) \), we let \( x(w) \in A_w \cap N_G(x(p_i)) \). We let \( W = V(P) \setminus V(T) \) and \( A = \bigcup_{w \in V(T)} (N_G[x(w)] \cup A_w) \).
Then $X, X_1, \ldots, X_s, Y_1, \ldots, Y_s$ are pairwise disjoint and pairwise anticomplete. Let $Z = X_1 \cup \cdots \cup X_s \cup Y_1 \cup \cdots \cup Y_s$, and let $G' = G|(Z \cup (V(G) \setminus A))$. It follows that

$$|N_{G'}(X_i)| \geq |N_G(X_i)| - |A| \geq C'\varepsilon n - |A| \geq C\varepsilon n,$$

where $C = C(|E(H) \setminus E(P)|, 2r)$. From our definition, it follows that $s = |E(H) \setminus E(P)|$. Moreover, since $v_w$ is an $r$-center for $A_w$ for all $w \neq v$, it follows that every vertex in $A_w$ is a $2r$-center for $A_w$ for all $w \neq v$ by Lemma 1.3.4.

By Lemma 3.2.10 it follows that there exist induced paths $P_1, \ldots, P_s$ in $G'$ such that

- for $i \in \{1, \ldots, s\}$, $P_i$ has one end in $X_i$ and one end in $Y_i$; and
- for $i, j \in \{1, \ldots, s\}, i \neq j$, $V(P_i)$ is anticomplete to $X_j \cup V(P_j) \cup Y_j$, and we can choose $P_i$ to be an $x_i$-$y_i$-path for all $i \in \{1, \ldots, s\}$.

For $i \in \{1, \ldots, s\}$, we have that $V(P_i) \subseteq A_{z_i^1} \cup A_{z_i^2} \cup (V(G) \setminus A)$. It follows that $V(P_i)$ is disjoint from $V(P')$, and $x(u_i), x(w_i)$ are the only vertices in $V(P')$ with a neighbor in $V(P_i)$. Let $P_i'$ be a shortest $x(u_i)$-$x(w_i)$-path in $(G|(V(P_i) \cup \{x(u_i), x(w_i)\})) \setminus x(u_i)x(w_i)$. It follows that $P_i'$ exists and has length at least two, and $P_i'$ is an induced $x(u_i)$-$x(w_i)$-path (possibly except for the edge $x(u_i)x(w_i)$), where $e_i = u_iw_i$. 77
Now let \( H' = G| (V(P') \cup V(P'_1) \cup \cdots \cup V(P'_s)) \). Then \( H' \) is isomorphic to a \( P \)-subdivision of \( H \): every edge of \( P \) is present un subdivided in \( G| V(P') \), and for every edge \( e = e_i = u_iw_i \in E(H) \setminus E(P) \), \( P'_i \) is an \( x(u_i)-x(w_i) \)-path with interior anticomplete to the remaining vertices of \( P' \). This contradiction implies the result. □

A \((k,c)\)-safe cover in a graph \( G \) with \(|V(G)| = n\) consists of disjoint sets \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \) with the following properties:

- \( A_i \) covers \( B_i \), \( B_i \) covers \( C_i \), and \( A_i \) is anticomplete to \( C_i \) for all \( i = 1, \ldots, k \);
- \( G|A_i \) is connected and \( A_i \) is anticomplete to \( A_j \cup B_j \cup C_j \) for all \( i, j = 1, \ldots, k \) with \( i \neq j \);
- \( B_i \) is anticomplete to \( C_j \) for all \( i, j = 1, \ldots, k \) with \( i < j \) and
- \( |C_i| \geq cn \) for all \( i = 1, \ldots, k \).

For a \((k,c)\)-safe cover \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \), we let its union \((A, B, C)\) be defined as \( A = A_1 \cup \cdots \cup A_k, B = B_1 \cup \cdots \cup B_k, C = C_1 \cup \cdots \cup C_k \).

We first prove that safe covers exist.

**Lemma 3.2.20.** Let \( C > 0 \) and \( k \in \mathbb{N} \). Then there exists an \( \varepsilon > 0 \) such that for every \( \varepsilon \)-coherent graph \( G \), \( G \) contains a \((k,C\varepsilon)\)-safe cover \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \) with \(|A_i| \leq \varepsilon n \) and \(|B_i| \leq 2\varepsilon n \) for all \( i = 1, \ldots, k \).

**Proof.** Let \( \varepsilon = 1/(2(4 + C + 2k)k) \). We will prove that given an \((i, (C + 2(k - i))\varepsilon)\)-safe cover \( A_1, \ldots, A_i, B_1, \ldots, B_i, C_1, \ldots, C_i \) and a set \( D \) with the following additional properties:

- \(|D| \geq (1 - (4 + C + 2k)i\varepsilon)n|;
- \( D \) is disjoint from \( A_1, \ldots, A_i, B_1, \ldots, B_i, C_1, \ldots, C_i \); and
- \((N(B_j) \cup N(A_j)) \cap D = \emptyset \) for all \( j \neq i \).
we can construct an \((i + 1, (C + 2(k - i - 1))\varepsilon)\)-safe cover

\[A_1, \ldots, A_{i+1}, B_1, \ldots, B_{i+1}, C'_1, \ldots, C'_{i+1}\]

and a set \(D'\) with the following additional properties:

- \(|D'| \geq (1 - (4 + C + 2k)(i + 1)\varepsilon)n;\)

- \(D'\) is disjoint from \(A_1, \ldots, A_{i+1}, B_1, \ldots, B_{i+1}, C'_1, \ldots, C'_{i+1}\); and

- \((N(B_j) \cup N(A_j)) \cap D' = \emptyset\) for all \(j \in \{1, \ldots, i + 1\} \).

To prove this, let \(A_1, \ldots, A_i, B_1, \ldots, B_i, C_1, \ldots, C_i\) and \(D\) be as above. Then \(|D| \geq n/2 \geq 3\varepsilon n,\) and so by Lemma 3.2.7, it follows that \(D\) contains a connected component \(D'\) with \(|D'| \geq 2\varepsilon n\). It follows that \(|N_G[D']| \geq (1 - \varepsilon)n\). Now, by Lemma 1.3.3 we let \(D' = \{d_1, \ldots, d_s\}\) such that \(G| \{d_1, \ldots, d_t\}\) is connected for all \(t \in \{1, \ldots, s\}\).

We let \(A_{i+1} = \{d_1, \ldots, d_t\}\) where \(t\) is minimum subject to \(|N_G[A_{i+1}]| \geq \varepsilon n\). Let \(B'_{i+1} = N_G(A_{i+1})\). Since \(|N_G[A_{i+1}]| \geq \varepsilon n\), it follows that \(|N_G[B'_{i+1}]| \geq (1 - \varepsilon)n\). By the minimality of \(t\), it follows that \(|N_G[A_{i+1}]| \leq 2\varepsilon n\), and so \(|B'_{i+1}| \leq 2\varepsilon n\), and since \(|N_G[A_{i+1}]| < (1 - \varepsilon)n\), it follows that \(|A_{i+1}| \leq \varepsilon n\). It follows that

\[
|N(B'_{i+1}) \cap (D \setminus (A_{i+1} \cup B'_{i+1}))| \geq |D| - |A_{i+1}| - |B'_{i+1}| - \varepsilon n
\]

\[
\geq (1 - (4 + C + 2k)i\varepsilon - 3\varepsilon)n
\]

\[
\geq n/2 \geq (C + 2(k - i))\varepsilon n.
\]

We let \(B_{i+1} \subseteq B'_{i+1}\) be minimal with respect to inclusion subject to \(|N(B'_{i+1}) \cap (D \setminus (A_{i+1} \cup B'_{i+1}))| \geq (C + 2(k - i - 1))\varepsilon n\), and let \(C'_{i+1} = N(B'_{i+1}) \cap (D \setminus (A_{i+1} \cup B'_{i+1}))\).

For \(j \in \{1, \ldots, i\}\), we let \(C'_j = C_j \setminus B'_{i+1}\). We let \(D' = D \setminus (A_{i+1} \cup B'_{i+1} \cup C_{i+1})\).

It follows that \(G|A_{i+1}\) is connected, \(A_{i+1}\) covers \(B_{i+1}\), \(B_{i+1}\) covers \(C'_{i+1}\), and \(A_{i+1}\) is
anticomplete to \( C'_{i+1} \). Moreover,

\[
|D'| \geq |D| - |A_{i+1} \cup B_{i+1} \cup C_{i+1}|
\]
\[
\geq |D| - (3\varepsilon + (C + 2(k - i - 1))\varepsilon + \varepsilon)n
\]
\[
\geq (1 - (4 + C + 2k)i\varepsilon - 4\varepsilon - (C + 2k)\varepsilon)n
\]
\[
\geq (1 - (4 + C + 2k)(i + 1)\varepsilon)n,
\]

and by definition, \( D' \) is disjoint from \( A_1, \ldots, A_{i+1}, B_1, \ldots, B_{i+1}, C'_1, \ldots, C'_{i+1} \). Since \( B'_{i+1} = N(A_{i+1}) \) and \( C_{i+1} = N(B_{i+1}) \cap D \), and since \( D' \subseteq D \), it follows that \( (N(B_j) \cup N(A_j)) \cap D' = \emptyset \) for all \( j \in \{1, \ldots, i + 1\} \). Since \( |B'_{i+1}| \leq 2\varepsilon \), it follows that

\[
|C'_j| \geq |C_j| - 2\varepsilon n \geq ((C + 2(k - i)) - 2)\varepsilon n = (C + 2(k - i - 1))\varepsilon n
\]

for all \( j \in \{1, \ldots, i\} \). Since \( C'_{i+1} \subseteq D \), it follows that \( B_j \) is anticomplete to \( C'_{i+1} \) for all \( j \in \{1, \ldots, i\} \). Since \( A_{i+1} \subseteq D \), it follows that \( A_{i+1} \) is anticomplete to \( A_j \cup B_j \) for all \( j \in \{1, \ldots, i\} \), and since \( N(A_{i+1}) = B'_{i+1} \) is disjoint from \( C'_j \) for all \( j \in \{1, \ldots, i\} \), it follows that \( A_{i+1} \) is anticomplete to \( C'_j \). Since \( C'_{i+1} \subseteq D \), it follows that \( A_j \) is anticomplete to \( C'_j \) for all \( j \in \{1, \ldots, i\} \). Finally, since \( B_{i+1} \subseteq D \cup C_1 \cup \cdots \cup C_i \), it follows that \( A_1 \cup \cdots \cup A_i \) is anticomplete to \( B_{i+1} \). This proves our claim.

We now let \( D_0 = V(G) \), and apply the claim \( k \) times; let

\[
A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k
\]

denote the resulting \((k, C\varepsilon)\)-safe cover. Then this satisfies the required properties, and the result follows. \( \square \)
A pairing \( M \) of a set \( X \) is a partition of \( X \) into sets of size at most two. If \( G \) is a graph, and \( X \subseteq V(G) \), then a collection of induced paths \( P_1, \ldots, P_k \) in \( G \) achieves \( M \) if:

- \( V(P_i) \) is anticomplete to \( V(P_j) \) for all \( i \neq j \) with \( i, j \in \{1, \ldots, k\} \);
- for every \( \{x, y\} \in M \), there exists an \( i \in \{1, \ldots, k\} \) such that the ends of \( P_i \) are \( x \) and \( y \); and
- for every \( \{x\} \in M \), \( x \) is anticomplete to \( V(P_1) \cup \cdots \cup V(P_k) \).

The following lemma shows that safe covers are useful for achieving pairings.

**Lemma 3.2.21.** Let \( k \in \mathbb{N} \). Then there exists an \( \varepsilon > 0 \) such that if \( G \) is \((\varepsilon, 2)\)-coherent, \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \) is a \((k, 3k\varepsilon)\)-safe cover in \( G \) with union \((A, B, C)\), and \( x_1 \in B_1, \ldots, x_k \in B_k \) are pairwise non-adjacent. Then for every pairing \( M \) of \( X = \{x_1, \ldots, x_k\} \), there exists a collection \( \mathcal{P} \) of paths contained in \( A \cup B \cup C \), and \( \mathcal{P} \) achieves \( M \).

**Proof.** Let \( D = \bigcup_{x \in X} N^2[x] \), and let \( C'_i = C_i \setminus D \) for \( i \in \{1, \ldots, k\} \). Since \( G \) is \((\varepsilon, 2)\)-coherent, it follows that \(|D| \leq k\varepsilon \), and so \( C'_i \subseteq |C_i| - k\varepsilon \geq 3k\varepsilon - k\varepsilon = 2k\varepsilon \). We let \( B'_i = B_i - N[X] \) for all \( i \in \{1, \ldots, k\} \). Since \( C'_i \cap N^2[X] = \emptyset \), it follows that \( B'_i \) covers \( C'_i \), and so \( A_1, \ldots, A_k, B'_1, \ldots, B'_k, C'_1, \ldots, C'_k \) is a \((k, 2k\varepsilon)\)-safe cover with union \((A, B', C')\) such that \( N[X] \cap (B' \cup C') = \emptyset \), and such that for all \( i, j \in \{1, \ldots, k\} \), \( x_i \) has a neighbor in \( A_j \) if and only if \( i = j \).

Let \( \{y_1, y'_1\}, \ldots, \{y_l, y'_l\} \) be the sets of size two in \( M \). It follows that \( l \leq k/2 \). We will prove by induction on \( k \) that if \( A_1, \ldots, A_k, B'_1, \ldots, B'_k, C'_1, \ldots, C'_k \) is a \((k, 2k\varepsilon)\)-safe cover with union \((A, B', C')\) such that \( N[X] \cap (B' \cup C') = \emptyset \) and \( x_1 \in B'_1, \ldots, x_k \in B'_k \) are pairwise non-adjacent, then for every pairing \( M \) of \( X = \{x_1, \ldots, x_k\} \), there exists a collection of paths that achieves \( M \).
If \( k = 0 \), or if \( M \) contains a set \( \{x_i\} \) for some \( i \in \{1, \ldots, k\} \), then the result follows by induction applied to \( A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_k, B'_1, \ldots, B'_{i-1}, B'_{i+1}, \ldots, B'_k, C'_1, \ldots, C'_{i-1}, C'_{i+1}, \ldots, C'_k \) and \( M \setminus \{\{x_i\}\} \).

Therefore, we may assume that \( M \) contains a set \( \{x_i, x_j\} \). Since \( |C'_i|, |C'_j| \geq 2k\varepsilon n \geq 2\varepsilon n \), it follows that there exists an edge \( ab \in E(G) \) with \( a \in C'_i, b \in C'_j \). Let \( a' \in B'_i \cap N(a), b' \in B'_j \cap N(b) \). Let \( P_a \) be an \( a'-x_i \)-path with \( V(P_a) \setminus \{a', x_i\} \subseteq A_i \); such a path exists since \( A_i \) is connected and \( a', x_i \) both have a neighbor in \( A_i \). Similarly, we let \( P_b \) be a \( b'-x_j \)-path with \( V(P_b) \setminus \{b', x_j\} \subseteq A_j \). It follows that \( G|(V(P_a) \cup V(P_b) \cup \{a, b\}) \) contains an induced \( x_i-x_j \)-path \( P \) such that \( V(P) \setminus \{a, b, a', b', x_i, x_j\} \subseteq A_i \cup A_j \).

We now modify \( B'_s, C'_s \) for all \( s \in \{1, \ldots, k\} \setminus \{i, j\} \) as follows: We let \( B''_s = B'_s \setminus N(a) \cup N(b) \cup N(a') \cup N(b') \), and we let \( C''_s = C_s \setminus N^2[\{a, b, a', b'\}] \). It follows that \( B''_s \) covers \( C''_s \), and

\[
|C''_s| \geq |C'_s| - 4\varepsilon n \geq 2k\varepsilon n - 4\varepsilon n = 2(k-2)\varepsilon n.
\]

Now we have obtained a \((k, 2(k-2))\)-safe cover, and \( N[V(P)] \subseteq A_i \cup B'_i \cup C'_i \cup A_j \cup B'_j \cup C'_j \). We remove \( A_i, B'_i, C'_i, A_j, B'_j, C'_j \) from the \((k, 2(k-2)\varepsilon)\)-safe cover; then \( N[V(P)] \) is disjoint from each set in the union of the resulting \((k-2, 2(k-2)\varepsilon)\)-safe cover. The result now follows by induction applied to this new cover, \( X \setminus \{x_i, x_j\} \) and \( M \setminus \{\{x_i, x_j\}\} \).

\[\square\]

**Lemma 3.2.22.** Let \( H \) be a graph, and let \( P \) be a (not necessarily induced) path in \( H \). Then there exists an \( r \geq 1 \) and \( \varepsilon > 0 \) such that every \((\varepsilon, r)\)-coherent graph contains a \( P \)-subdivision of \( H \).

**Proof.** Let \( T \) be a caterpillar with spine \( P \), head \( v \) (where \( v \) is an end of \( P \)), and in which every vertex \( p \in V(P) \) has degree \( d_H(p) \). We define a map \( f : V(T) \setminus V(P) \rightarrow E(H) \setminus E(P) \) such that for every edge \( e \in E(H) \setminus E(P) \), \( f^{-1}(e) = \{x, y\} \), where
Let \( v \) and \( w \) be the neighbor of \( \varepsilon \) in \( \varepsilon \). We choose \( v \) and \( w \) be the neighbor of \( \varepsilon \) in \( \varepsilon \). Now let \( c, k, C' \) be as above. By Lemma 3.2.20, it follows that for every \( c, k, C' \)-safe cover \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \) with \( |A_i| \leq \varepsilon n \) and \( |B_i| \leq 2\varepsilon n \) for all \( i \in \{1, \ldots, k\} \). Fix such an \( \varepsilon > 0 \), and let \( G \) be \((\varepsilon, r)\)-coherent. Let \( n = |V(G)| \).

Let \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \) be a \((k, (C' + D + 1)\varepsilon)\)-safe cover. We let \( C'_i \subseteq C_i \) be minimum with respect to inclusion subject to \( |C_i| \geq C'\varepsilon n \) for all \( i \in \{1, \ldots, k\} \), and we let \( D_i = C_i \setminus C'_i \) for all \( i \in \{1, \ldots, k\} \). It follows that \( |D_i| \geq D\varepsilon n \) for all \( i \in \{1, \ldots, k\} \). Now \( C = C'_1, \ldots, C'_k \) is a \((k, C'\varepsilon)\)-box array. Thus, by Lemma 3.2.14, it follows that there is a \( C \)-rainbow \((T, v)\)-cover structure \( \{A_t\}_{t \in V(T)} \) with \( |A_v| \geq C\varepsilon n \) and such that \( A_t \) is connected and \( |N_G(A_t)| \geq cn \) for all \( t \in V(T) \).

We now apply Lemma 3.2.11 to obtain a function \( x : V(T) \to V(G) \) such that \( G|_{\{x(V(T))\}} \) is isomorphic to \( T \) with isomorphism \( x^{-1} \), and \( x(t) \in A_t \) for all \( t \in V(T) \). Let \( v_1, \ldots, v_m \) be an ordering of \( V(T) \setminus V(P) \) such that for \( i \in \{1, \ldots, m\} \), \( v_i \in C'_{j_i} \) and \( j_1 < \cdots < j_m \). By symmetry, we may assume that \( j_i = i \) for all \( i \in \{1, \ldots, m\} \).

We choose \( x_i \in B_i \) and a path \( Q_i \) for \( i \in \{1, \ldots, m\} \), as follows: Since \( |N(A_v)| \geq 2\varepsilon n \) and \( G \) is \((\varepsilon, r)\)-controlled, it follows that there is a vertex \( y \in N(A_v) \) with \( d(y, x(v)) \geq r + 1 \). It follows that there is a vertex \( z \in A_v \) with \( d(z, x(v)) = r \geq |V(T)| + 4m \). Let \( w \) be the neighbor of \( v_i \) in \( V(T) \). Since \( G|_{A_v} \) is connected, we

By Lemma 3.2.14, it follows that for every \( C > 0 \), there exist \( c, k, C' > 0 \) with \( c \geq C\varepsilon \) such that for every \((k, C'\varepsilon)\)-box array \( B \), and for every \( \varepsilon \)-coherent graph \( G \), \( G \) contains a \( B \)-rainbow \((T, v)\)-cover structure \( \{A_t\}_{t \in V(T)} \) with \( |A_v| \geq cn \), where \( n = |V(G)| \), and by choosing \( A_t \) according to Lemma 3.2.12 applied to a \((k', c')\)-box array with \( c' \geq c \), we may assume that \( A_t \) is connected and \( |N_G(A_t)| \geq cn \) for all \( t \in V(T) \).

Now let \( C = 2 \), \( D = 3m + 1 \), and let \( c, k, C' \) be as above. By Lemma 3.2.20, it follows that there exists an \( \varepsilon > 0 \) such that for every \( \varepsilon \)-coherent graph \( G \), \( G \) contains a \((k, (C' + D + 1)\varepsilon)\)-safe cover \( A_1, \ldots, A_k, B_1, \ldots, B_k, C_1, \ldots, C_k \) with \( |A_i| \leq \varepsilon n \) and \( |B_i| \leq 2\varepsilon n \) for all \( i \in \{1, \ldots, k\} \). Fix such an \( \varepsilon > 0 \), and let \( G \) be \((\varepsilon, r)\)-coherent. Let \( n = |V(G)| \).
may choose a shortest \( x(w) \)-z-path \( P \) in \( G \mid (A_{v_i} \cup \{x(w)\}) \). By the properties of a \((T, v)\)-cover structure, we may assume that \( x(v_i) \) is the neighbor of \( x(w) \) in \( V(P) \).

Let \( y_i \) be the vertex closest to \( x(w) \) along \( P \) that satisfies \( d(y_i, x(v)) \geq |V(T)| + 4i \). Let \( x_i \) be a neighbor of \( y_i \) in \( B_i \), and let \( Q_i = x(v_i)Py_i-x_i \).

\[
\text{For all } i \in \{1, \ldots, m\}, V(Q_i) \setminus \{x(v_i)\} \text{ is anticomplete to } x(V(T) \setminus \{v_i\}) \tag{3.3}
\]

and anticomplete to \( V(Q_j) \) for all \( j \in \{1, \ldots, m\} \setminus \{i\} \).

Let \( w \) be the neighbor of \( v_i \) in \( V(T) \). Since \( V(Q_i) \setminus \{x_i\} \subseteq A_{v_i} \), it follows that \( A_t \) is anticomplete to \( V(Q_i) \setminus \{x_i\} \) for all \( t \in V(T) \setminus \{w\} \). By construction, \( N(x(w)) \cap V(Q_i) = \{x_i\} \). Now suppose that \( x_i \) has a neighbor in \( V(T) \). By construction, \( d(x_i, x(v)) \geq |V(T)| + 3 \), and so \( d(x_i, x(y)) \geq 3 \) for all \( y \in V(T) \). This proves the first half of the claim.

For the second half, let \( i, j \in \{1, \ldots, m\} \) with \( i < j \). Then \( V(Q_i) \) is anticomplete to \( V(Q_j) \setminus \{x_j\} \), since \( B_i \) is anticomplete to \( C_j \supseteq (V(Q_j) \setminus \{x_j\}) \), and \( A_{v_i} \) is anticomplete to \( A_{v_j} \supseteq (V(Q_j) \setminus \{x_j\}) \), and since \( V(Q_i) \subseteq B_i \cup A_{v_i} \). Since \( d(x_j, x(v)) \geq |V(T)| + 4j - 1 \) and since, by the choice of \( Q_i \), \( d(y, x(v)) \leq |V(T)| + 4i + 1 \) for all \( y \in V(Q_i) \), it follows that \( d(x_j, y) \geq |V(T)| + 4j - 1 - |V(T)| - 4i - 1 \geq 2 \) for all \( y \in V(Q_i) \) since \( j > i \). This implies (3.3).

Let \( Y = x(V(T)) \cup \bigcup_{i \in \{1, \ldots, m\}} V(Q_i) \). It follows that \( N^2[Y] \subseteq N^r[x(v)] \), since \( d(y, x(v)) \leq |V(T)| + 5m + 1 = r - 2 \) for all \( y \in Y \). Consequently, \( |N^2[Y]| \leq \varepsilon n \).

We let \( B'_i = B_i \setminus (N[Y] \setminus \{x_i\}) \) and \( D'_i = D_i \setminus N^2[Y] \) for \( i \in \{1, \ldots, m\} \). It follows that \( B'_i \) covers \( D'_i \) and \( |D'_i| \geq 3m \varepsilon n \). Let \( A_1, \ldots, A_m, B'_1, \ldots, B'_m, D'_1, \ldots, D'_m \) be the \((m, 3m \varepsilon)\)-safe cover we obtain from this. Let \( M \) be the pairing that contains \( \{x_i, x_j\} \) for all \( e \in E(H) \) such that \( f^{-1}(e) = \{v_i, v_j\} \). Now, by Lemma 3.2.21, it follows that there is a collection \( P_1, \ldots, P_{m/2} \) that achieves this pairing. Now \( J = G \mid (Y \cup V(P_1) \cup \cdots \cup V(P_{m/2})) \) is a \( P \)-subdivision of \( H \): Since \( P \) is the spine...
of \( T \), we choose \( y : V(P) \to V(T) \) be the map that takes a vertex in \( P \) to the corresponding vertex in \( T \). It follows that \( x(y(V(P))) \) is a path in \( G|Y \). For every edge \( e = ab \in E(H) \setminus E(P) \), there is a pair \( \{x_i, x_j\} \in M \) where \( x(a) = u, x(b) = w \), and \( v_i \in N_T(u), v_j \in N_T(w) \). Therefore, there is a path \( P_s \) that has ends \( x_i, x_j \), and \( x(w)-x(v_i)Q_i x_i P_s x_j Q_j x(v_j)-x(u) \) is the path in \( J \) that corresponds to \( e \). This shows that \( J \) is a \( P \)-subdivision of \( H \), and thus the result is proved.

We are now ready to prove the main result of this section.

**Proof of Theorem 3.2.18.** By Lemma 3.2.19 it follows that for all \( \varepsilon > 0, r \geq 1 \), there exists a \( c > 0 \) such that if there is a \( c \)-coherent graph that does not contain an induced \( P \)-subdivision of \( H \), then there is a \( (\varepsilon, r) \)-coherent graph that does not contain an induced \( P \)-subdivision of \( H \). By Lemma 3.2.22, there exist \( \varepsilon > 0, r \geq 1 \), such that this is impossible; and thus there is a \( c > 0 \) such that no graph that does not contain an induced \( P \)-subdivision of \( H \) is \( c \)-coherent. This implies Theorem 3.2.18.

A tree \( T \) is a **subdivided caterpillar** if there is a path \( P \) in \( T \) such that every vertex of degree at least three in \( T \) is contained in \( V(P) \). The path \( P \) is called a **spine** for \( T \).

**Theorem 3.2.23.** Let \( T \) be a subdivided caterpillar. Then there exists an \( \varepsilon > 0 \) such that every \( T \)-free \( \varepsilon \)-sparse graph \( G \) contains an anticomplete \( (\varepsilon n, \varepsilon n) \)-pair, where \( n = |V(G)| \).

**Proof.** Let \( P \) be a spine for \( T \). Then every \( P \)-subdivision of \( T \) contains \( T \) as an induced subgraph. Therefore, every \( T \)-free graph \( G \) also does not contain a \( P \)-subdivision of \( T \) as an induced subgraph. Therefore the result follows from Theorem 3.2.18.

### 3.2.5 All trees

In this section, we prove that for every tree \( T \), there exists an \( \varepsilon > 0 \) such that every \( \varepsilon \)-coherent graph contains \( T \) as an induced subgraph. Our main proof strategy is
based on rainbow matchings, defined below; we first consider the general case and show how excluding a tree implies the presence of big rainbow matchings, and we treat the case in which a big rainbow matching exists at the end of this section.

Let \(B_1, \ldots, B_k\) be a \((k, c)\)-box array. For \(k' \leq k\) and \(c' \leq c\), a \((k', c')\)-box array \(A_1, \ldots, A_{k'}\) is \textit{contained in} \(B_1, \ldots, B_k\) if there is a function \(g : \{1, \ldots, k'\} \rightarrow \{1, \ldots, k\}\) such that \(g(i) \neq g(j)\) for all \(i, j \in \{1, \ldots, k'\}\) with \(i \neq j\), and \(A_i \subseteq B_{f(i)}\) for all \(i \in \{1, \ldots, k'\}\). If additionally it is true that \(g(i) < g(j)\) for all \(i, j \in \{1, \ldots, k'\}\) with \(i < j\), we say that \(A_1, \ldots, A_{k'}\) is \textit{order-contained in} \(B_1, \ldots, B_k\).

A \((k', c')\)-\textit{rainbow matching} in a \((k, c)\)-box array \(B_1, \ldots, B_k\) is a \((k', c')\)-box array \(B'_1, \ldots, B'_{k'}\) and a collection of \textit{cover sets} \(A_1, \ldots, A_{k'}\) such that

- for all \(i \in \{1, \ldots, k'\}\), \(A_i\) covers \(B_i\), and \(A_i\) is anticomplete to \(B_j\) for all \(j \in \{1, \ldots, k'\} \setminus \{i\}\);
- for all \(i \in \{1, \ldots, k'\}\), there exist \(j, j'\) such that \(A_i \subseteq B_j\) and \(B'_i \subseteq B_{j'}\);
- for all \(i \in \{1, \ldots, k\}\), there exists at most one \(j \in \{1, \ldots, k'\}\) such that \((A_j \cup B'_j) \cap B_i \neq \emptyset\); and
- for all \(i \in \{1, \ldots, k\}\), \(j \in \{1, \ldots, k'\}\), either \(A_j \cap B_i = \emptyset\) or \(B'_j \cap B_i = \emptyset\).

In particular, the \((k', c')\)-box array \(B'_1, \ldots, B'_{k'}\) is contained in \(B_1, \ldots, B_k\).

For a \((k, c)\)-box array \(B = B_1, \ldots, B_k\), and for \(f > 0\), we define the set of \textit{\(f\)-coverable sets} \(\mathcal{T}(B, f)\) as the set of all triples \((i, I, J)\) such that

- \(i \in \{1, \ldots, k\}\);
- \(I \cup J = \{1, \ldots, k\} \setminus \{i\}\);
- \(I\) and \(J\) are disjoint; and
- there exists a set \(A \subseteq B_i\) such that for every \(j \in \{1, \ldots, k\} \setminus \{i\}\), if \(j \in I\), then \(|B_j \cap N(A)| \geq cf n\); and if \(j \in J\), then \(|B_j \setminus N(A)| \geq cf n\).
We say that the set \( A \) (as in the last bullet) witnesses \((i, I, J)\).

It follows that \(|\mathcal{T}(B, f)| \leq k2^{k-1}\). Let \( B' = B'_1, \ldots, B'_k \) be a \((k, c')\)-box array order-contained in \( B_1, \ldots, B_k \) and let \( f' > 0 \) such that \( cf \leq c'f' \). It follows that \( \mathcal{T}(B', f') \subseteq \mathcal{T}(B, f) \), since for every \((i, I, J) \in \mathcal{T}(B', f')\), the set \( A \) that witnesses \((i, I, J)\) in \( \mathcal{T}(B', f') \) also witnesses \((i, I, J)\) in \( \mathcal{T}(B, f) \) since \( c'f' \geq cf \).

For \( t \geq 1 \), we say that \((B', f')\) \(t\)-improves \((B, f)\) if

- \( B' \) is a \((k, c')\)-box array order-contained in \( B \);
- \( f^{1/2^t} \geq f' > 0 \);
- \( |\mathcal{T}(B', f')| < |\mathcal{T}(B, f)| \); and
- \( c'f' \geq cf \).

If there is no \((B', f')\) that \(t\)-improves \((B, f)\), we say that \((B, f)\) is \(t\)-optimal. The following lemma demonstrates how to turn a \((k, c)\)-box array into a \(t\)-optimal \((k, c')\)-box array:

**Lemma 3.2.24.** Let \( G \) be a graph, let \( t, k \in \mathbb{N} \), and let \( c > 0 \). Let \( B = B_1, \ldots, B_k \) be a \((k, c)\)-box array in \( G \). Let \( f > 0 \). Then there exists a \((k, c')\)-box array \( B' \) order-contained in \( B \), and \( f' > 0 \) such that \((B', f')\) is \(t\)-optimal, \( f' \leq f^2 - tk2^{k-1} \), and \( c' \geq cf/f' \).

**Proof.** We define a sequence of box arrays \( B_0, \ldots, B_s \), and constants \( c_0, \ldots, c_s \), \( f_0, \ldots, f_s \), as follows: We let \( B_0 = B, c_0 = c, \) and \( f_0 = f \). For \( i \geq 0 \), \( B_i \) is a \((k, c_i)\)-box array, and if \( i > 0 \), then \((B_i, f_i)\) improves \((B_{i-1}, f_{i-1})\); and there is no \((B', f')\) that improves \((B_s, f_s)\). Since \(|\mathcal{T}(B, f)| \leq k2^{k-1}\), it follows that \( s \leq k2^{k-1} \), and therefore \( f_s \leq f_0^{2^t - tk2^{k-1}} \), and \( c_s \geq c_0f_0/f_s \), as claimed.

We need the following definition. Let \( G \) be a graph, and let \( k \in \mathbb{N} \) and \( c > 0 \). Let \( B = B_1, \ldots, B_k \) be a \((k, c)\)-box array. Let \( f > 0 \), and let \( \mathcal{T} = \mathcal{T}(B, f) \). A pseudo-matching of size \( t \) in \( \mathcal{T} \) consists of disjoint sets \( A, B \subseteq \{1, \ldots, k\} \) and a bijection.
and hence the lemma is proved.

The next lemma shows that pseudo-matchings in optimal box arrays give us rainbow matchings.

**Lemma 3.2.25.** Let $G$ be a graph, $t, k, c \in \mathbb{N}$, and let $c > 0$. Let $B = B_1, \ldots, B_k$ be a $(k, c)$-box array in $G$. Let $f > 0$ such that $(B, f)$ is $t$-optimal. If $T(B, f)$ has a pseudo-matching of size $t$, then there is a $(t, c')$-rainbow matching in $B$ with $c' = c\sqrt{t}$.

**Proof.** Let $A, B, h$ be a pseudo-matching of size $t$ in $T(B, f)$. By symmetry, we may assume that $A = \{1, \ldots, t\}$, $B = \{t + 1, \ldots, 2t\}$, and $h(i) = i + t$ for all $i \in A$.

For $i \in \{0, 1, \ldots, t\}$, we will inductively construct a $(k, c_i)$-box array $B'_1, \ldots, B'_t$ and cover sets $C_1, \ldots, C_i$ such that $B'_1, \ldots, B'_t, C_1, \ldots, C_i$ is an $(i, c_i)$-rainbow matching and such that $c_i \geq cf^{2^{i-t-1}}$ and $C_1, \ldots, C_i$ are anticomplete to $B'_{i+1}, \ldots, B'_t$.

For $i = 0$, $B$ is the desired box array. Now let $i \geq 0$, and suppose that we have constructed a $(k, c_i)$-box array $B' = B'_1, \ldots, B'_k$ and cover sets $C_1, \ldots, C_i$ as above. If $i = t$, then we have constructed a rainbow matching as desired. Now let $i < t$. Since $(B, f)$ is $t$-optimal, it follows that $(B', f')$ with $f' = f^{2^{i-t-1}}$ does not $t$-improve $(B, f)$. Since there is a triple $(i + 1, I, J) \in T(B, f)$ with $h(i + 1) \in I$ and $B \setminus \{h(i + 1)\} \subseteq J$, it follows that $(i + 1, I, J) \in T(B', f')$. Let $A \subseteq B'_{i+1}$ be a set that witnesses $(i + 1, I, J)$. It follows that $A$ has at least $c'f'n \geq cf^{2^{i-t-1}} \cdot f^{2^{i-t-1}}n = cf^{2^{i+1-t-1}}$ neighbors in $B'_{h(i+1)}$ and at least $c'f'n$ non-neighbors in each of $B'_s$ for $s \in B \setminus \{h(i + 1)\}$. We let $c_{i+1} = cf^{2^{i+1-t-1}}$, $C_{i+1} = A$, $B'_h = B'_h \cap N(A)$, $B'_s = B'_s \setminus N(A)$ for $s \in B \setminus \{h(i + 1)\}$, and $B''_s = B'_s$ for $s \in \{1, \ldots, k\} \setminus B$. It follows that $B''_1, \ldots, B''_{i+1}, C_1, \ldots, C_{i+1}$ is an $(i + 1, c_{i+1})$-rainbow matching, $c_{i+1} \geq cf^{2^{i+1-t-1}}$ and $C_1, \ldots, C_{i+1}$ are anticomplete to $B''_{i+2}, \ldots, B'_t$. This completes the inductive step, and hence the lemma is proved.
We now define a function \( h \) that captures the local covering patterns. Let \( G \) be a graph, \( k, p \geq 1, c > 0 \), and let \( \mathcal{B} = B_1, \ldots, B_k \) be a \((k, c)\)-box array. Let \( f > 0 \), and let \( \mathcal{T} = \mathcal{T}(\mathcal{B}, f) \). Let \( i_1 < \cdots < i_p \) with \( \mathcal{I} = \{i_1, \ldots, i_p\} \subseteq \{1, \ldots, k\} \). We let

\[
h(\mathcal{I}) = \mathcal{T}(B_{i_1}, \ldots, B_{i_p}, f).
\]

It follows that

\[
h(\mathcal{I}) = \{(j, \{l : i_l \in \mathcal{I} \cap I\}, \{l : i_l \in \mathcal{I} \cap J\}) : i_j \in \mathcal{I}, (i_j, I, J) \in \mathcal{T}(\mathcal{B}, f)\}.
\]

We call a \((k, c)\)-box array \((\mathcal{B}, f)\) \(p\)-homogeneous if \( h(\mathcal{I}) \) takes the same value for all \( \mathcal{I} \subseteq \{1, \ldots, k\} \) with \( |\mathcal{I}| = p \).

The next lemma shows how to make a box array homogeneous.

**Lemma 3.2.26.** Let \( p, s \in \mathbb{N} \). Then there exists a \( k = k(s, p) \in \mathbb{N} \) such that for all \( c > 0 \), every \((k, c)\)-box array \( \mathcal{B} \) order-contains a \(p\)-homogeneous \((s, c)\)-box array \( \mathcal{B}' \). Moreover, if \( \mathcal{B} \) is \( t \)-optimal, then \( \mathcal{B}' \) is \( t \)-optimal.

**Proof.** Let \( p, s \) be as in the statement of the lemma. We let \( k = R_p(c_1, \ldots, c_{p2^p-1}) \) with \( c_1 = \cdots = c_{p2^p-1} = s \) as in Theorem 1.3.1. Now let \( c > 0 \), and let \( \mathcal{B} \) be a \((k, c)\)-box array. For \( \mathcal{I} \subseteq \{1, \ldots, k\} \), the set \( h(\mathcal{I}) \) is a subset of the set of possible triples, which has size \( p2^{p-1} \), and consequently, there are at most \( 2^{p2^{p-1}} \) possible values of \( h \).

By Theorem 1.3.1 it follows that there exists a set \( \mathcal{J} \subseteq \{1, \ldots, k\} \) with \( |\mathcal{J}| = s \) such that \( h(\mathcal{I}) \) takes the same value for every \( \mathcal{I} \subseteq \mathcal{J} \) with \( |\mathcal{I}| = p \). But then \( \{B_j : j \in \mathcal{J}\} \) is \(p\)-homogeneous. The second statement follows from the definition of \(t\)-optimality.

Let \( \mathcal{B} = B_1, \ldots, B_k \) be a \((k, c)\)-box array and \( p \in \mathbb{N}, f > 0 \). Then \((\mathcal{B}, f)\) is \(p\)-interval if for every \((i, I, J) \in \mathcal{T}(\mathcal{B}, f)\), there exist \( j_1, j_2 \in \{1, \ldots, k\} \) such that for \( A = \{1, \ldots, j_1, j_2, \ldots, k\} \), we have \( I \subseteq A \), and \( |A \setminus I| \leq 2p \).

The next lemma shows that either we can get our box arrays to be \(p\)-interval, or we get a big pseudo-matching:
Lemma 3.2.27. Let \( t, k, p \in \mathbb{N} \) and let \( c > 0 \). Suppose that \( k \geq 8p \). Let \( G \) be a graph, and let \( B = B_1, \ldots, B_k \) be a \((k, c)\)-box array. Let \( f > 0 \) be such that \((B, f)\) is \((2p + 2)\)-homogeneous. Then either \( T(B, f) \) contains a pseudo-matching of size at least \( p \), or \((B, f)\) is \( p \)-interval.

Proof. Let \( t, k, p, c, f, G, B \) be as in the statement of the lemma. Suppose that \((B, f)\) is not \( p \)-interval. It follows that there exists a triple \((i, I, J) \in T(B, f)\) such that there do not exist \( j_1, j_2 \in \{1, \ldots, k\} \) such that for \( A = \{1, \ldots, j_1, j_2, \ldots, k\} \), we have \( I \subseteq A \), and \(|A \setminus I| \leq 2p\).

Suppose that for every \( j \in I \), either \(|\{j' < j : j' \in J\}| \leq p\) or \(|\{j' > j : j' \in J\}| \leq p\). Then we let \( j_1 \) be the maximum \( j \in \{1, \ldots, k\} \) for which the first condition holds, and let \( j_2 \) be the minimum \( j \in \{1, \ldots, k\} \) for which the second condition holds. By the monotonicity of the condition, it follows that every \( j \in I \), either \( j \leq j_1 \) or \( j \geq j_2 \); but then for \( A = \{1, \ldots, j_1, j_2, \ldots, k\} \), we have \( I \subseteq A \), and \(|A \setminus I| \leq 2p\). This proves that there is a \( j \in I \) such that \(|\{j' < j : j' \in J\}| > p\) and \(|\{j' > j : j' \in J\}| > p\).

Let \( I \) consist of \( \{i, j\} \) and \( p \) elements from each of the sets \( \{j' < j : j' \in J\} \), \( \{j' > j : j' \in J\} \). It follows that \( I \setminus \{i\} = \{i_1, \ldots, i_{2p+1}\} \) with \( i_1 < \cdots < i_{2p+1} \) and \( I \cap I = \{i_{p+1}\} \). Let \( r \in \{1, \ldots, 2p + 2\} \) be such that \( r = 1 \) if \( i < i_1 \); \( r = 2p + 2 \) if \( i > i_{2p+1} \); and \( i_{r-1} < i < i_r \) otherwise. By symmetry, we may assume that \( r > p + 1 \).

By the definition of the function \( h \) and the set \( I \), it follows that \( h(I) \) contains the triple \((r, \{p + 1\}, \{1, \ldots, 2p + 2\} \setminus \{r, p + 1\})\).

For all \( w \in \{1, \ldots, p\} \), we let \( I_w = \{2w, 2 + 2w, \ldots, 4p + 2w\} \cup \{2r - 3 + 2w\} \). Since \((B, f)\) is \((2p + 2)\)-homogeneous, it follows that for all \( w \in \{1, \ldots, p\} \), \( h(I_w) \) contains \((r, \{p + 1\}, \{1, \ldots, 2p + 2\} \setminus \{r, p + 1\})\). Let \( A_w \) be a set that witnesses \((r, \{p + 1\}, \{1, \ldots, 2p + 2\} \setminus \{r, p + 1\})\) in \( h(I_w) \). Then \( A_w \) witnesses a triple \((2r - 3 + 2w, I_w, J_w)\) in \((B, f)\) that corresponds to \((r, \{p + 1\}, \{1, \ldots, 2p + 2\} \setminus \{r, p + 1\})\) in \( h(I_w) \). Let \( A = \{2r - 3 + 2w : r \in \{1, \ldots, p\}\} \), and let \( B = \{2p + 2, 2p + 4, \ldots, 4p\} \), and let \( g : A \to B \) be a function defined by setting \( g(2r - 3 + 2w) = 2w + 2p \). It
follows that $g(2r - 3 + 2w) \in I_w$, since $A_w$ witnesses $(r, \{p + 1\}, \{1, \ldots, 2p + 2\} \setminus \{r, p + 1\})$ in $h(I_w)$, and $2w + 2p$ is bigger than $p$ elements of $I_w$ and less than $p + 1$ elements. By construction, we have that $B \subseteq I_w$ for all $w \in \{1, \ldots, p\}$, and therefore $B \setminus \{g(2r - 3 + 2w)\} \subseteq J_w$. This implies that $A, B, g$ define a pseudo-matching of size $p$ in $T(B, f)$. \hfill \square

The following lemma gives sufficient conditions for preserving the box array properties of being optimal and being homogeneous when restricting to a box array contained in an optimal and homogeneous box array.

**Lemma 3.2.28.** Let $G$ be a graph, $t, k \geq 1$, $c \geq 0$, and let $B = B_1, \ldots, B_k$ be a $(k, c)$-box array in $G$. Let $f > 0$, and suppose that $(B, f)$ is $t$-optimal. Let $B'$ be a $(k', c')$-box array order-contained in $B$, and let $f' > 0$ and $t' < t$ such that $f' \leq f^{1/2'}$ and $c'f' \geq cf$. Then $(B', f')$ is $(t - t')$-optimal.

Furthermore, if $(B, f)$ is $p$-homogeneous, then $(B', f')$ is $p$-homogeneous; and if $(B, f)$ is $p'$-interval, then $(B', f')$ is $p'$-interval.

**Proof.** Suppose not; and let $f'', c'' > 0$, and let $(B'', f'')$ be a $(k', c'')$-box array that $(t - t')$-improves $(B', f')$. Let $B^*$ arise from $B'' = B_1'', \ldots, B_k''$, by letting $B_i^* = B_j''$ if there is a $j \in \{1, \ldots, k'\}$ such that $B_j'' \subseteq B_i$, and by letting $B_i^* = B_i$ otherwise.

We have that $c''f'' \geq c'f' \geq cf$ and $f'' \geq f'^{1/2(c' - c)} \geq f^{1/2t}$. Let $(i, I, J) \in T(B', f') \setminus T(B'', f'')$, and let $A \subseteq B_i'$ witness $(i, I, J)$. It follows that there is a triple $(i', I', J')$ witnessed by $A$ in $(B, f)$. Now suppose that there is a set $A'$ witnessing $(i', I', J')$ in $(B^*, f'')$. Then $A'$ witnesses $(i, I, J)$ in $(B', f')$, a contradiction. It follows that $(i', I', J') \in T(B, f) \setminus T(B^*, f'')$. This proves that $B^*$ $t$-improves $B$, a contradiction, since $B$ is $t$-optimal.

For the second part of the lemma, note that since $(B', f')$ does not $t'$-improve $(B, f)$, it follows that $T(B', f')$ arises from $T(B, f)$ by restricting to those $j \in \{1, \ldots, k\}$ for which there is a $j' \in \{1, \ldots, k'\}$ with $B_j \cap B_{j'} \neq \emptyset$, and by mapping $j$ to $j'$ in this
way. This implies that for \((B', f')\), \(h(I)\) takes the same value for all \(I \subseteq \{1, \ldots, k'\}\) with \(|I| = p\).

For the third statement, we let \(B' = B'_1, \ldots, B'_k\) and let \(A \subseteq B'_i\) witness a triple \((i, I, J) \in \mathcal{T}(B', f')\). It follows that \(A\) witnesses a triple \((i', I', J') \in \mathcal{T}(B, f)\). Since \((B, f)\) is \(p'\)-interval, it follows that there exist \(j_1, j_2 \in \{1, \ldots, k\}\) such that \(I' \subseteq \{1, \ldots, j_1, j_2, \ldots, k\}\) and \(|\{1, \ldots, j_1, j_2, \ldots, k\} \setminus I'| \leq 2p'\). Now let \(j'_1\) be the largest \(j\) such that there is an \(i \in \{1, \ldots, j_1\}\) with \(B'_i \subseteq B_i\), and let \(j'_2\) be the smallest \(j\) such that there is an \(i \in \{j_2, \ldots, k\}\) with \(B'_i \subseteq B_i\). It follows that \(I \subseteq \{1, \ldots, j'_1, j'_2, \ldots, k'\}\), and that \(|\{1, \ldots, j'_1, j'_2, \ldots, k'\} \setminus I| \leq |\{1, \ldots, j_1, j_2, \ldots, k\} \setminus I'| \leq 2p'\). This implies that \((B', f')\) is \(p'\)-interval.

We now give some definitions related to trees and the ways in which they are contained in box arrays. Recall that \(T_{d,h}\) is defined as a tree \(T\) with a root \(v \in V(T)\) such that \(d_T(v) = d\), and \(d_T(w) \in \{1, d + 1\}\) for all \(w \in V(T) \setminus \{v\}\), and such that \(d_T(v, w) = h\) for every leaf \(w\) of \(T\).

We define \(T_{d,h}(i)\) as the tree \(T\) obtained from \(i + 1\) copies of \(T_{d,h}\), say \(T_1, \ldots, T_{i+1}\) with roots \(v_1, \ldots, v_{i+1}\), by adding an edge \(v_jv_{i+1}\) for all \(j \in \{1, \ldots, i\}\). We call \(v_{i+1}\) the root of \(T_{d,h}(i)\), and we call \(T_1, \ldots, T_i\) big branches of \(T_{d,h}(i)\). In particular, \(T_{d,h}(0)\) is isomorphic to \(T_{d,h}\) and the root of \(T_{d,h}(0)\) is the root of \(T_{d,h}\). Moreover, \(T_{d,h}(d)\) is isomorphic to \(T_{d,h+1}\).

Let \(B = B_1, \ldots, B_k\) be a box array in a graph \(G\). Let \(H\) be an induced subgraph of \(G\). We say that \(H\) is \(B\)-rainbow if \(|V(H) \cap B_i| \leq 1\) for all \(i \in \{1, \ldots, k\}\), and \(V(H) \subseteq B_1 \cup \cdots \cup B_k\).

Now let \(T = T_{d,h}(i)\) be a tree and let \(v\) be its root. Suppose that there is an induced subgraph \(H\) of \(G\) such that \(H\) is \(B\)-rainbow and isomorphic to \(T\). Let \(g : V(T) \rightarrow V(H)\) be an isomorphism. We say that \(H\) is a left-rooted \(T\) if \(g(v) \in B_i\) for some \(i \in \{1, \ldots, k\}\), and \(g(w) \in B_j\) with \(j > i\) for all \(w \in V(T) \setminus \{v\}\). We say
that $H$ is a right-rooted $T$ if $g(v) \in B_i$ for some $i \in \{1, \ldots, k\}$, and $g(w) \in B_j$ with $j > i$ for all $w \in V(T) \setminus \{v\}$.

Let $d, h \in \mathbb{N}$. Let $\mathcal{S}$ denote the set of all trees $T_{d,h'}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d - 1\}$. For a $(k, c)$-box array $\mathcal{B}$, we let $\mathcal{L}(\mathcal{B}, \mathcal{S})$ denote the set of all pairs $(I, T)$ satisfying that

- $I \subseteq \{1, \ldots, k\}$, $I = \{i_1, \ldots, i_l\}$ with $i_1 < \cdots < i_l$;
- $T \in \mathcal{S}$ and $l = |V(T)|$; and
- $\mathcal{B}' = B_{i_1}, \ldots, B_{i_l}$ contains a left-rooted $T$.

Similarly, we let $\mathcal{R}(\mathcal{B}, \mathcal{S})$ denote the set of all pairs $(I, T)$ satisfying that

- $I \subseteq \{1, \ldots, k\}$, $I = \{i_1, \ldots, i_l\}$ with $i_1 < \cdots < i_l$;
- $T \in \mathcal{S}$ and $l = |V(T)|$; and
- $\mathcal{B}' = B_{i_1}, \ldots, B_{i_l}$ contains a right-rooted $T$.

If $\mathcal{B}'$ is a $(k, c')$-box array order-contained in $\mathcal{B}$, then $\mathcal{L}(\mathcal{B}', \mathcal{S}) \subseteq \mathcal{L}(\mathcal{B}, \mathcal{S})$ and $\mathcal{R}(\mathcal{B}', \mathcal{S}) \subseteq \mathcal{R}(\mathcal{B}, \mathcal{S})$ by the definition above. For $f > 0$, a $(k, c')$-box array $\mathcal{B}'$ is said to $(\mathcal{S}, f)$-improve $\mathcal{B}$ if

- $\mathcal{B}'$ is order-contained in $\mathcal{B}$ and $c' \geq cf$; and
- $|\mathcal{L}(\mathcal{B}', \mathcal{S})| + |\mathcal{R}(\mathcal{B}', \mathcal{S})| < |\mathcal{L}(\mathcal{B}, \mathcal{S})| + |\mathcal{R}(\mathcal{B}, \mathcal{S})|$.

If there is no $(k, c')$-box array that $(\mathcal{S}, f)$-improves $\mathcal{B}$, then $\mathcal{B}$ is $(\mathcal{S}, f)$-optimal.

**Lemma 3.2.29.** Let $d, h \in \mathbb{N}$. Let $\mathcal{S}$ denote the set of all trees $T_{d,h'}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d - 1\}$. Let $M = 2|V(T_{d,h})|$. Let $k \in \mathbb{N}, c, f > 0$, and let $\mathcal{B}$ be a $(k, c)$-box array. Then there is a $(k, c')$-box array $\mathcal{B}'$ such that $\mathcal{B}'$ is order-contained in $\mathcal{B}$, $\mathcal{B}'$ is $(\mathcal{S}, f)$-optimal, and $c' \geq cf^{2(h+1)dkM/2}$.  

93
Proof. We define a sequence of box arrays $\mathcal{B}_0 = \mathcal{B}, \mathcal{B}_1, \ldots, \mathcal{B}_s$ and numbers $c_0 = c, c_1, \ldots, c_s$ as follows: Suppose that $\mathcal{B}_i$ has been defined, and is a $(k, c_i)$-box array. If there is a $(k, c')$-box array $\mathcal{B}'$ that $(\mathcal{S}, \mathcal{f})$-improves $\mathcal{B}_i$, then we let $\mathcal{B}_{i+1} = \mathcal{B}'$ and $c_{i+1} = c' \geq cf$. Otherwise we stop and let $s = i$.

It follows that $\mathcal{B}_s$ is $(\mathcal{S}, \mathcal{f})$-optimal. Since $|\mathcal{S}| \leq (h' + 1)d$, it follows that $|L(\mathcal{B}, \mathcal{S})| + |R(\mathcal{B}, \mathcal{S})| \leq 2(h' + 1)dk^{M/2}$, and therefore $s \leq 2(h' + 1)dk^{M/2}$. It follows that $c_s \geq cf^s \geq c^2f^{2(h' + 1)dk^{M/2}}$, as claimed. This proves the lemma.

Let $d, h \in \mathbb{N}$. Let $\mathcal{S}$ denote the set of all trees $T_{d,h'}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d - 1\}$. A $(k, c)$-box array $\mathcal{B}$ is $\mathcal{S}$-smooth if for every $T \in \mathcal{S}$, the following two statements hold:

- either $L(\mathcal{B}, \mathcal{S})$ contains no pair $(I, T)$ with $I \subseteq \{1, \ldots, k\}$; or for every $I \subseteq \{1, \ldots, k\}$ with $|I| = |V(T)|$, $(I, T) \in L(\mathcal{B}, \mathcal{S})$; and
- either $R(\mathcal{B}, \mathcal{S})$ contains no pair $(I, T)$ with $I \subseteq \{1, \ldots, k\}$; or for every $I \subseteq \{1, \ldots, k\}$ with $|I| = |V(T)|$, $(I, T) \in R(\mathcal{B}, \mathcal{S})$.

If $\mathcal{B}$ is $\mathcal{S}$-smooth, we call $T \in \mathcal{S}$ left-good for $\mathcal{B}$ if for every $I \subseteq \{1, \ldots, k\}$ with $|I| = |V(T)|$, $(I, T) \in L(\mathcal{B}, \mathcal{S})$; and right-good for $\mathcal{B}$ if for every $I \subseteq \{1, \ldots, k\}$ with $|I| = |V(T)|$, $(I, T) \in R(\mathcal{B}, \mathcal{S})$.

The next lemma shows how to turn a box array into a smooth box array.

Lemma 3.2.30. Let $d, h \in \mathbb{N}$. Let $\mathcal{S}$ denote the set of all trees $T_{d,h'}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d - 1\}$. Let $k \in \mathbb{N}$. There is a $K = K(k, d, h) \in \mathbb{N}$ such that for all $c > 0$, every $(K, c)$-box array $\mathcal{B}$ order-contains an $\mathcal{S}$-smooth $(k, c)$-box array $\mathcal{B}'$.

Proof. It it sufficient to prove that for every $T \in \mathcal{S}$ and $k \in \mathbb{N}$, there exists a $K \in \mathbb{N}$ such that every $(K, c)$-box array $\mathcal{B}$ order-contains a $(k, c)$-box array $\mathcal{B}'$ such that either $L(\mathcal{B}', \mathcal{S})$ contains no pair $(I, T)$ with $I \subseteq \{1, \ldots, k\}$; or for every $I \subseteq \{1, \ldots, k\}$ with |
|I| = |V(T)|, (I, T) ∈ L(𝔉′, ℳ). The result of the lemma then follows by symmetry between left and right, and by applying this claim for every \( T \in ℳ \) once for \( L(𝔉, ℳ) \) and once for \( ℛ(𝔉, ℳ) \).

Let \( K = R_{|V(T)|}(k, k) \). Let \( 𝔉 \) be a \((K, c)\)-box array, and for every set \( S \subseteq \{1, \ldots, K\} \) with \( |S| = |V(T)| \), we let \( s(S) = 1 \) if \((I, T) \in L(𝔉, ℳ)\) and \( s(S) = 0 \) otherwise. By Theorem 1.3.1 there exists a set \( \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, K\} \) with \( i_1 < \cdots < i_k \) such that \( s(S) \) takes the same value for all \( S \subseteq \{i_1, \ldots, i_k\} \). This proves our claim by letting \( 𝔉' = 𝔉_{i_1}, \ldots, 𝔉_{i_k} \), and the statement of the lemma follows.

The next lemma gives sufficient conditions for maintaining smoothness and \((ℳ, f)\)-optimality.

**Lemma 3.2.31.** Let \( d, h \in \mathbb{N} \). Let \( ℳ \) denote the set of all trees \( T_{d,h'}(i) \) for \( h' \in \{0, \ldots, h\} \) and \( i \in \{0, \ldots, d-1\} \). Let \( k \in \mathbb{N}, c, f > 0 \), and let \( 𝔉 \) be a \((k, c)\)-box array. Let \( 𝔉' \) be a \((k', c')\)-box array \( 𝔉' \) such that \( 𝔉' \) is order-contained in \( 𝔉 \). Let \( 1 \geq f' > 0 \) such that \( c' f' \geq cf \). Then \( 𝔉' \) is \((ℳ, f')\)-optimal. Furthermore, if \( 𝔉 \) is \( ℳ\)-smooth, then so is \( 𝔉' \).

**Proof.** We begin by proving the first statement of the lemma. Suppose that it does not hold; i.e. suppose that there is a \((k'', c'')\)-box array \( 𝔉'' = 𝔉''_1, \ldots, 𝔉''_k'' \) that \((ℳ, f')\)-improves \( 𝔉' \). Let \( 𝔇^* \) be a \((k, c'')\)-box array with \( B^*_i = B''_j \) if there is a \( j \in \{1, \ldots, k''\} \) such that \( B''_j \subseteq B_i \), and with \( B^*_i = B_i \) otherwise, for all \( i \in \{1, \ldots, k\} \). Since \( c'' \geq c' f' \geq cf \), and since \( 𝔇^* \) does not \((ℳ, f)\)-improve \( 𝔉 \), it follows that

\[
|L(𝔇^*, ℳ)| + |R(𝔇^*, ℳ)| = |L(𝔉, ℳ)| + |R(𝔉, ℳ)|,
\]

and consequently, \( L(𝔇^*, ℳ) = L(𝔉, ℳ) \) and \( R(𝔇^*, ℳ) = R(𝔉, ℳ) \). Since \( 𝔉'' \) is an \((ℳ, f')\)-improvement of \( 𝔉' \), by symmetry, we may assume that there is a pair \((I, T) \in L(𝔉', ℳ) \setminus L(𝔉'', ℳ) \). Let \( I' = \{j \in \{1, \ldots, k\} : \exists i \in I : B'_i \subseteq B_j\} \). It follows that
\((I', T) \in \mathcal{L}(\mathcal{B}, \mathcal{S})\), and so \((I', T) \in \mathcal{L}(\mathcal{B}^*, \mathcal{S})\). But for all \(i \in I', B_i^* = B_i''\), and so 
\((I, T) \in \mathcal{L}(\mathcal{B}'', \mathcal{S})\), a contradiction. This proves the first statement of the lemma.

We are now ready to prove the second statement of the lemma. Suppose that we have \(I, J, T\) with \(|I| = |J| = |V(T)|\) such that \((I, T) \in \mathcal{L}(\mathcal{B}', \mathcal{S})\) and \((J, T) \notin \mathcal{L}(\mathcal{B}', \mathcal{S})\). Let \(I' = \{j \in \{1, \ldots, k\} : \exists i \in I : B_i' \subseteq B_j\}\) and \(J' = \{j \in \{1, \ldots, k\} : \exists i \in J : B_i' \subseteq B_j\}\). It follows that \((I', T) \in \mathcal{L}(\mathcal{B}, \mathcal{S})\), and therefore, since \(\mathcal{B}\) is \(\mathcal{S}\)-smooth, it follows that \((J', T) \in \mathcal{L}(\mathcal{B}, \mathcal{S})\). Now let \(\mathcal{B}^*\) be a \((k, c')\)-box array with \(B_i^* = B_j'\) if there is a \(j \in \{1, \ldots, k'\}\) such that \(B_j' \subseteq B_i\), and with \(B_i^* = B_i\) otherwise, for all \(i \in \{1, \ldots, k\}\). It follows that \((J', T) \notin \mathcal{L}(\mathcal{B}, \mathcal{S})\). Since \(c' \geq cf\), it follows that \(\mathcal{B}'\) is an \((\mathcal{S}, f)\)-improvement for \(\mathcal{B}\), contrary to the \((\mathcal{S}, f)\)-optimality of \(\mathcal{B}\). This is a contradiction, and the second statement of the lemma follows.

\[
\text{Lemma 3.2.32. Let } d, h, c, f \in \mathbb{N}. \text{ Let } \mathcal{S} \text{ denote the set of all trees } T_{d,h}(i) \text{ for } h' \in \{0, \ldots, h\} \text{ and } i \in \{0, \ldots, d-1\}. \text{ Let } k \in \mathbb{N} \text{ and } c, f > 0. \text{ Let } \mathcal{B} \text{ be a } (k, c)\text{-box array} \text{ such that } \mathcal{B} \text{ is } \mathcal{S}\text{-smooth and } (\mathcal{S}, f)\text{-optimal. Let } a, \alpha \in \mathbb{N} \text{ with } \alpha < d, \text{ and suppose that}
\]
\( k = |V(T_{d,a}(\alpha))| \) and \( T_{d,a}(\alpha) \) is left-good. Then there is a left \((T_{d,a}(\alpha), f)\)-partition of \( B \). If instead we assume that \( T_{d,a}(\alpha) \) is right-good, then there is a right \((T_{d,a}(\alpha), f)\)-partition of \( B \).

**Proof.** By symmetry, it suffices to prove the lemma in the case when \( T_{d,a}(\alpha) \) is left-good. Let \( V_1, \ldots, V_s \) be a collection of disjoint sets such that for every \( i \in \{1, \ldots, s\} \), \( G|V_i \) is a left-rooted \( T_{d,a}(\alpha) \), and suppose that this collection is maximal, i.e. there is no \( V_{s+1} \) disjoint from \( V_1 \cup \cdots \cup V_s \) such that \( G|V_{s+1} \) is a left-rooted \( T_{d,a}(\alpha) \).

Let \( V = \bigcup_{i \in \{1, \ldots, s\}} V_i \) and let \( B' = B_1 \setminus V, \ldots, B_k \setminus V \). It follows that \( (\{1, \ldots, k\}, T_{d,a}(\alpha)) \notin L(B', S) \). Since \( B \) is \((S, f)\)-optimal, it follows that \( B' \) does not \((S, f)\)-improve \( B \). Since \( |L(B', S)| < |L(B, S)| \), this implies that there is no \( c' \geq cf \) such that \( B' \) is a \((k, c')\)-box array. It follows that there is an \( i \in \{1, \ldots, k\} \) such that \( |B_i \setminus V| < cf n \), and therefore, \( |B_i \cap V| \geq (1-f)cn \). By construction, it follows that \( |V \cap B_j| = s \) for all \( j \in \{1, \ldots, k\} \), and therefore \( s \geq (1-f)cn \). This implies that \( V_1, \ldots, V_s \) is a left \((T_{d,a}(\alpha), f)\)-partition.

Let \( T = T_{d,h}(i) \) be a tree with root \( v \) for some \( d, h, i \in \mathbb{N} \). Let \( B \) be a box array in a graph \( G \) and let \( A, B \subseteq V(G) \) be disjoint. Then \( A \) is said to \( T \)-cover \( B \) if

- \( A = V(T_1) \cup \cdots \cup V(T_s) \) where \( G|T_i \) is a left-rooted or a right-rooted \( T \) for all \( i \in \{1, \ldots, s\} \); and

- for all \( b \in B \), there exists an \( i \in \{1, \ldots, s\} \) such that \( b \) is adjacent to the vertex of \( T_i \) corresponding to \( v \) in the isomorphism between \( T \) and \( G|T_i \), and non-adjacent to all other vertices in \( T_i \).

It is a left \( T \)-cover if \( G|T_i \) is always a left-rooted \( T \) in the first bullet; and a right \( T \)-cover if \( G|T_i \) is always a right-rooted \( T \) for all \( i \in \{1, \ldots, s\} \).

This next lemma shows how to build \( T \)-covers covering one of the first or last \( 2p \) boxes in \( p \)-interval box arrays:
Lemma 3.2.33. Let $d, h \in \mathbb{N}$. Let $S$ denote the set of all trees $T_{d,h}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d-1\}$.

Let $k, p \in \mathbb{N}$ and $f, c > 0$ with $f \leq 1/2$. Let $\varepsilon \leq c/4$ and let $G$ be $\varepsilon$-coherent. Let $\mathcal{B} = B_1, \ldots, B_k$ be a $(k, c)$-box array in $G$ such that $\mathcal{B}$ is $S$-smooth, $(S, f)$-optimal, and $(\mathcal{B}, f)$ is $p$-interval. Let $a, \alpha, b, \beta \in \mathbb{N}$ with $\alpha < d$ and $\alpha, b < h$, and suppose that $M_1 = |V(T_{d,a}(\alpha))|$, $M_2 = |V(T_{d,b}(\beta))|$ and $T_{d,a}(\alpha)$ is left-good, $T_{d,b}(\beta)$ is right-good.

Let $\mu \in \mathbb{N}$, and suppose that $f \leq 1/(16\mu M^2)$ and $f \leq \frac{\mu - 1}{6(4p+1)\mu}$. Let $m \geq \frac{6(4p+1)\mu}{\mu - 1}$.

Let $R = \{1, \ldots, 2p, \ldots, k - 2p + 1, \ldots, k\}$. Suppose that $k \geq 2p(M_1 + M_2) + 4p$, and let $M = M_1 + M_2$. Suppose further that $\varepsilon \leq \frac{\mu - 1}{6(4p+1)\mu}c$.

Then there is a set $A$ such that $A \subseteq \bigcup_{i \in \{2p+1, \ldots, 2p(M_1 + M_2 + 1)\}} B_i$, and a $(k, c/m)$-box array $\mathcal{B}'$ order-contained in $\mathcal{B}$ such that:

- $\mathcal{B}' = B'_1, \ldots, B'_k$;
- there exists a $j \in R$ such that $A$ is either a left $T_{d,a}(\alpha)$-cover or a right $T_{d,b}(\beta)$-cover of $B'_j \subseteq B_j$; and
- for all $j^* \in \{1, \ldots, k\} \setminus \{j, 2p + 1, 2p + 2, \ldots, 2p(M + 1)\}$, the set $A$ is anticomplete to $B'_{j^*}$.

Proof. Let $G, d, h, S, k, p, f, c, \varepsilon, a, \alpha, b, \beta, m, \mathcal{B}, \mu, M_1, M_2$ and $M$ be as in the statement of the lemma.

For $j \in \{1, \ldots, 2p\}$, we define sets $I_j, J_j$ and $l(j), r(j)$ as follows:

- $I_j = \{2p + j\} \cup \{4p + (M_1 - 1)(j - 1) + 1, \ldots, 4p + (M_1 - 1)j\}$; $l(j) = 2p + j$; and
- $J_j = \{\alpha + (j - 1)(M_2 - 1) + 1, \ldots, \alpha + j(M_2 - 1)\} \cup \{2pM + j\}$; where $\alpha = 2p(M_1 + 1)$, and $r(j) = 2pM + j$.

By definition, we have that $l(j) \leq i \leq r(j')$ for all $j, j', j'' \in \{1, \ldots, 2p\}$ and $i \in I_{j''} \cup J_{j''}$.
For \( j \in \{1, \ldots, 2p\} \), we let \( I_j = \{i^1_j, \ldots, i^{M_1}_j\} \) with \( i^1_j < \cdots < i^{M_1}_j \). By Lemma 3.2.32, it follows that there is a left \((T_{d,a}(\alpha), f)\)-partition \( V^j_1, \ldots, V^j_{s_j} \) of \( B_{i^1_j}, \ldots, B_{i^{M_1}_j} \).

For \( j \in \{1, \ldots, 2p\} \), we let \( J_j = \{i^1_j, \ldots, i^{M_2}_j\} \) with \( i^1_j < \cdots < i^{M_2}_j \). By Lemma 3.2.32, it follows that there is a right \((T_{d,b}(\beta), f)\)-partition \( U^j_1, \ldots, U^j_{t_j} \) of \( B_{i^1_j}, \ldots, B_{i^{M_2}_j} \).

We let

\[
\mathcal{R} = \bigcup_{j \in \{1, \ldots, 2p\}} \left( \bigcup_{i \in \{1, \ldots, s_j\}} (V^j_i \cap B_{i(j)}) \cup \bigcup_{i \in \{1, \ldots, t_j\}} (U^j_i \cap B_{r(j)}) \right).
\]

By construction, it follows that for every \( j \in \{1, \ldots, 2p\} \), and for every \( i \in \{1, \ldots, s_j\} \), \( \mathcal{R} \cap V^j_i = \{r\} \) where \( r \) is the root of the left-rooted \( T_{d,a}(\alpha) \) given by \( G|V^j_i \); and for every \( i \in \{1, \ldots, t_j\} \), \( \mathcal{R} \cap U^j_i = \{r\} \) where \( r \) is the root of the right-rooted \( T_{d,b}(\alpha) \) given by \( G|U^j_i \). We let

\[
\mathcal{C} = \left( \bigcup_{j \in \{1, \ldots, 2p\}} \left( \bigcup_{i \in \{1, \ldots, s_j\}} V^j_i \cup \bigcup_{i \in \{1, \ldots, t_j\}} U^j_i \right) \right) \setminus \mathcal{R}.
\]

Let \( T \) be an induced subgraph of \( G \) isomorphic to one of the trees \( T_{d,a}(\alpha), T_{d,b}(\beta) \), and let \( r \) be the vertex of \( T \) corresponding to the root of the tree. Let \( v \in V(G) \setminus V(T) \). We say that \( v \) hits \( T \) well if \( N(v) \cap V(T) = \{r\} \); \( v \) hits \( T \) badly if \( N(v) \cap (V(T) \setminus \{r\}) \neq \emptyset \); and \( v \) misses \( T \) if \( N(v) \cap V(T) = \emptyset \).

Let \( B^* = \bigcup_{i \in \mathcal{R}} B_i \). We construct a sequence \( e_1, \ldots, e_s \) of vertices and sequences \( T_0, \ldots, T_s, I^*_0, \ldots, I^*_s, J^*_0, \ldots, J^*_s \) of sets, with the following properties:

- \( I^*_0 = J^*_0 = T_0 = \emptyset \);
- for all \( i \in \{1, \ldots, s\} \), \( e_i \in B^* \setminus \{e_1, \ldots, e_{i-1}\} \);
- for all \( i \in \{1, \ldots, s\} \), \( \{e_1, \ldots, e_{i-1}\} \) is anticomplete to \( (\mathcal{R} \cup \mathcal{C}) \setminus T_i \);
- for all \( i \in \{1, \ldots, s\} \), \( T_i = \bigcup_{(j,j') \in I^*_i} V^j_{j'} \cup \bigcup_{(j,j') \in J^*_i} U^j_{j'} \);
- for all \( i \in \{1, \ldots, s\} \), \( T_{i-1} \subseteq T_i \);
• for all \( i \in \{1, \ldots, s\} \), and for all \((j, j') \in I_i^* \setminus I_{i-1}^*\), \(e_i\) has a neighbor in \(V_j^2\);

• for all \( i \in \{1, \ldots, s\} \), and for all \((j, j') \in J_i^* \setminus J_{i-1}^*\), \(e_i\) has a neighbor in \(U_j^2\);

• for all \( i \in \{1, \ldots, s\} \), let

\[
B_i^L = \{(j, j') \in I_i^* \setminus I_{i-1}^*: e_i \text{ hits } G|V_j^2 \text{ badly}\}
\]

\[
B_i^R = \{(j, j') \in J_i^* \setminus J_{i-1}^*: e_i \text{ hits } G|U_j^2 \text{ badly}\};
\]

then \(\mu(|B_i^L| + |B_i^R|) \geq |I_i^* \setminus I_{i-1}^*| + |J_i^* \setminus J_{i-1}^*|\); and

• for every vertex \( v \in B^* \setminus \{e_1, \ldots, e_s\} \),

\[
\mu\left|\{(j, j'): (j, j') \not\in I_i^*; v \text{ hits } G|V_j^2 \text{ badly}\}\right|
\]

\[
+ \mu\left|\{(j, j'): (j, j') \not\in J_i^*; v \text{ hits } G|U_j^2 \text{ badly}\}\right|
\]

\[
\leq \left|\{(j, j'): (j, j') \not\in I_i^*; N(v) \cap V_j^2 \neq \emptyset\}\right|
\]

\[
+ \left|\{(j, j'): (j, j') \not\in J_i^*; N(v) \cap U_j^2 \neq \emptyset\}\right|.
\]

To construct these sequences, suppose that \(e_1, \ldots, e_{i-1}, I_{i-1}^*, J_{i-1}^*\) have been defined and satisfy all but the last bullet. We may assume that the condition in the last bullet does not hold for \(s = i - 1\), otherwise we let \(s = i - 1\) and stop. Let \(v \in B^* \setminus \{e_1, \ldots, e_{i-1}\}\) be such that the last bullet does not hold for \(v\) and \(s = i - 1\).

We let \(e_i = v\), and

\[
I_i^* = I_{i-1}^* \cup \{(j, j'): N(v) \cap V_j^2 \neq \emptyset\}; \quad J_i^* = J_{i-1}^* \cup \{(j, j'): N(v) \cap U_j^2 \neq \emptyset\}.
\]

We let \(T_i = \bigcup_{(j, j') \in T_i} V_j^2 \cup \bigcup_{(j, j') \in J_i} U_j^2\). By the choice of \(v\), it follows that \(e_1, \ldots, e_i, T_i, I_i^*, J_i^*\) satisfy all bullets but the last.
We claim that this process does not collect too many of our trees:

\begin{align}
(I.4) & \text{ When this process terminates, we have } |I_s^*| + |J_s^*| \leq \frac{\varepsilon}{2} n.
\end{align}

Suppose not. Let \( E = \{e_1, \ldots, e_i\} \) with \( i \in \{1, \ldots, s\} \) minimum such that \(|I_s^*| + |J_s^*| > \frac{\varepsilon}{2} n\). Since \( \mu(|B^L_j| + |B^R_j|) \geq |I_j^* \setminus I_{j-1}| + |J_j^* \setminus J_{j-1}| \) for all \( j \in \{1, \ldots, i\} \), it follows that

\[
|N(E) \cap C| \geq \sum_{j=1}^{i} \mu(|B^L_j| + |B^R_j|) \geq \frac{1}{\mu} (|I_j^*| + |J_j^*|) > \frac{c}{2} n \geq 8p^2(M_1 + M_2)cfn.
\]

Since \( E \subseteq B^* \), it follows that there is a \( j^* \in R \) such that \( E' = E \cap B_{j^*} \) satisfies

\[
|N(E) \cap C| \geq 2p(M_1 + M_2)cfn.
\]

It follows that there is a \( j \in \{2p + 1, \ldots, 2p(M_1 + M_2 + 1)\} \) such that \( |N(E') \cap B_j| \geq cfn \). Furthermore, \( |N(E') \cap \mathcal{R}| \leq |I_j| + |J_j| \leq \frac{\varepsilon}{2} n \leq (1 - f)cn \), and therefore, \( |B_{j'} \setminus N(E')| \geq cfn \) for all \( j' \in \{l(1), \ldots, l(2p), r(1), \ldots, r(2p)\} \). But now \( E' \) witnesses the triple \((j^*, I, J) \in \mathcal{T}(B, f)\) with \( j \in I \) and \( \{l(1), \ldots, l(2p), r(1), \ldots, r(2p)\} \subseteq J \). Since \( l(1) \leq \cdots \leq l(2p) \leq j \leq r(1) \leq \cdots \leq r(2p) \), it follows that \((B, f)\) is not \( p \)-interval, a contradiction. This proves (3.4).

We let \( e_1, \ldots, e_s, \mathcal{T}_s, I_s^*, J_s^* \) be as defined above, and let \( E = \{e_1, \ldots, e_s\} \). Let

\[
\Gamma = \{V^j_{j'} : (j, j') \notin I_s^*\} \cup \{U^j_{j'} : (j, j') \notin J_s^*\}
\]

be the set of trees whose vertex sets are not included in \( \mathcal{T}_s \). Since \(|I_s^* + J_s^*| < \frac{\varepsilon}{2} n\), it follows that \( |\Gamma| \geq \frac{c(1-2f)}{2} n \). It follows that for every \( v \in B^* \setminus E \), we have

\[
\mu|\{W \in \Gamma : v \text{ hits } G|W \text{ badly}\}| \leq |\{W \in \Gamma : N(v) \cap W \neq \emptyset\}|.
\]

Therefore, if we order the trees in \( \Gamma \) randomly, it follows that the probability that the first tree \( W \) such that \( v \) does not miss \( W \) is a tree that \( v \) hits well, is at least
Let $W_1, \ldots, W_\gamma$ be such that $\Gamma = \{W_1, \ldots, W_\gamma\}$, with the ordering chosen at random. We say that $v$ is \textit{proper} for this ordering of $\Gamma$ if the first tree $W$ that $v$ does not miss is hit well by $v$.

It follows that for all $r \in R$, the expected number of vertices in $B_r$ that are proper for $W_1, \ldots, W_\gamma$ is at least $|B_r|(\mu - 1)/\mu$. The probability that the number of proper vertices in $B_r$ is less than $|B_r|(\mu - 1)/(4p + 1)\mu$ is at most $\frac{1}{4p+1}$. By the union bound, and since $|R| = 4p$, it follows that there is an ordering $W_1, \ldots, W_\gamma$ of $\Gamma$ such that for all $r \in R$, the set $B^*_r$ of vertices in $B_r$ that are proper for $W_1, \ldots, W_\gamma$ is at least $c\mu - 1/2(4p+1)\mu n$.

It follows that $|B^*_r| \geq 3\varepsilon n$ for all $r \in R$. Since $\gamma \geq \frac{c-2f}{2}n$, and $f \leq \frac{1}{4}$, and $\frac{c}{4} \geq \varepsilon$, it follows that $|W_1 \cup \cdots \cup W_\gamma| \geq \varepsilon n$. Since $G$ is $\varepsilon$-coherent, it follows that $|B^*_r \setminus N(W_1 \cup \cdots \cup W_\gamma)| \leq \varepsilon n$ for all $r \in R$. Now let $s \in \{1, \ldots, \gamma\}$ be minimum such that there exists an $r \in R$ with $|N(W_1 \cup \cdots \cup W_s) \cap B^*_r| \geq c\frac{\mu - 1}{2(4p+1)\mu} n$. By the minimality of $s$, and since $G$ is $\varepsilon$-coherent, it follows that for all $r' \in R$,

$$|B^*_r \setminus N(W_1 \cup \cdots \cup W_s)| > c\frac{\mu - 1}{2(4p+1)\mu} n - \varepsilon n \geq c\frac{\mu - 1}{6(4p+1)\mu} n \geq cf n.$$  

Since $(\mathcal{B}, f)$ is $p$-interval, it follows that for all $j \in \{2p+1, \ldots, 2p(M+1)\}$, we have that $|B_q \cap N((W_1 \cup \cdots \cup W_s) \cap B_j)| \leq cf n$ for all $q \in \{1, \ldots, k\} \setminus (R \cup \{2p+1, \ldots, 2p(M+1)\})$. Therefore, $|B_q \cap N(W_1 \cup \cdots \cup W_s)| \leq c(2pM) fn \leq cn/2$ for all $q \in \{1, \ldots, k\} \setminus (R \cup \{2p+1, \ldots, 2p(M+1)\})$.

Now let

$$A_1 = \bigcup_{i \in \{1, \ldots, s\}, G|W_i \text{ is a left-rooted } T_{d,a}(\alpha)} W_i$$

and

$$A_2 = \bigcup_{i \in \{1, \ldots, s\}, G|W_i \text{ is a right-rooted } T_{d,b}(\beta)} W_i.$$
Let \( \eta \in \{1, 2\} \) be such that the number of vertices in \( B_\eta^* \) that are hit well by \( W_i \) for some \( i \in \{1, \ldots, s\} \) with \( W_i \subseteq A_\eta \) is at least \( c \frac{\mu - 1}{6(4p+1)\mu} n \). Such \( \eta \) exists, since for every vertex \( v \) in \( N(W_1 \cup \cdots \cup W_s) \cap B_\eta^* \), either \( A_1 \) or \( A_2 \) contains a tree \( W_i \) with \( i \in \{1, \ldots, s\} \) such that \( v \) hits \( G|W_i \) well. Now let \( A = A_\eta \), and let \( B'_r \) consist of the vertices in \( B_\eta^* \) that are hit well by \( W_i \) for some \( i \in \{1, \ldots, s\} \) with \( W_i \subseteq A_\eta \). Let \( B'_r = B_r \setminus N[A] \) for all \( r' \in \{1, \ldots, k\} \setminus \{2p + 1, \ldots, 2p(M + 1)\} \), and let \( B'_r = B_r \) for all \( r \in \{2p + 1, \ldots, 2p(M + 1)\} \). It follows that \( |B'_r| \geq c \frac{\mu - 1}{6(4p+1)\mu} n \geq cn/m \) for all \( r' \in \{1, \ldots, k\} \). Moreover, \( A \) either left \( T_{d,a}(\alpha) \)-covers \( B'_r \) (if \( \eta = 1 \)) or \( A \) right \( T_{d,b}(\beta) \)-covers \( B'_r \) (if \( \eta = 2 \)). Finally, by construction, \( A \) is anticomplete to \( B'_r \) for all \( r' \in \{1, \ldots, k\} \setminus \{r, 2p + 1, 2p + 2, \ldots, 2p(M + 1)\} \). This proves that \( A \) has the desired properties, and completes the proof.

In the following lemma, we show how to turn a collection of \( T \)-covers into bigger and better trees.

**Lemma 3.2.34.** Let \( d, h \in \mathbb{N} \). Let \( S \) denote the set of all trees \( T_{d,h'}(i) \) for \( h' \in \{0, \ldots, h\} \) and \( i \in \{0, \ldots, d - 1\} \).

Let \( k, p, f, c > 0 \) with \( f \leq 1/2 \). Let \( \varepsilon \leq c/4 \) and let \( G \) be \( \varepsilon \)-coherent. Let \( B \) be a \( (k, c) \)-box array in \( G \) such that \( B \) is \( S \)-smooth, \( (S', f) \)-optimal, and \( (B', f) \) is 1-optimal and \( p \)-interval. Let \( a, \alpha, b, \beta \in \mathbb{N} \) with \( \alpha, \beta < d \) and \( a \leq b < h \), and let \( M_1 = |V(T_{d,a}(\alpha))|, M_2 = |V(T_{d,b}(\beta))| \).

Suppose that \( T_{d,a}(\alpha) \) is left-good, and that \( T_{d,b}(\beta) \) is right-good. We let \( M = M_1 + M_2 \).

Suppose that

- \( k \geq 16p^2dM + 4p \);
- \( f \leq (1/(60Mp^2)) \cdot (60p)^{-8pd} \cdot (60p)^{-8pd} \); and
- \( \varepsilon \leq 1/(60p)^{8pd+1} \cdot c \).
Then either $T_{d,a+1}$, $T_{d,b+1}$ or $T_{d,a}(\alpha + 1)$ is left-good, or $T_{d,b+1}$ is right-good.

Proof. Since $\mathcal{B}$ is $\mathcal{S}$-smooth, it suffices to prove that $\mathcal{B}$ contains either a left-rooted $T_{d,a+1}$, a left-rooted $T_{d,b+1}$, a left-rooted $T_{d,a}(\alpha + 1)$ or a right-rooted $T_{d,b+1}$. We let $\mu = 2$ and $m = 60p$; let $R = \{1, \ldots, 2p, k - 2p + 1, \ldots, k\}$.

There is a sequence of box arrays $\mathcal{B}_0 = \mathcal{B}, \mathcal{B}_1, \ldots, \mathcal{B}_{8pd}$, sets $A_1, \ldots, A_{8pd}$, and numbers $f_0, \ldots, f_{8pd}$ such that the following hold for all $i \in \{1, \ldots, 8pd\}$:

- $\mathcal{B}_i$ is a $(k, c/m^i)$-box array order-contained in $\mathcal{B}$;
- $f_i \geq f \cdot m^i$;
- $\mathcal{B}_i$ is $\mathcal{S}$-smooth, $(\mathcal{S}, f_i)$-optimal, and $(\mathcal{B}, f_i)$ is $p$-interval;
- $\mathcal{B}_i = B^i_1, \ldots, B^i_k$;
- for every $j \in \{1, \ldots, i\}$, there exists a $j' = j'(i, j) \in R$ such that $A_j$ either $T_{d,a}(\alpha)$-covers or $T_{d,b}(\beta)$-covers $B^i_{j'} \subseteq B_{j'}$;
- for every $j \in \{1, \ldots, i\}$, $A_j$ is anticomplete to $B_{j^*}$ for $j^* \in \{1, \ldots, k\} \setminus (\{j'\} \cup \{2p + 1, \ldots, 2p(iM + 1)\})$, where $j' = j'(i, j)$ as in the previous bullet;
- for every $j \in \{1, \ldots, i - 1\}$, $A_j$ is anticomplete to $A_i$; and
- $A_i \subseteq B_{2p(M(i-1)+1)+1}, \ldots, B_{2p(iM+1)}$.

We prove this by induction on $i \in \{0, \ldots, 8pd\}$. For $i = 0$, these statements hold. Now let $i > 0$, and suppose that $\mathcal{B}_{i-1}$ and $A_1, \ldots, A_{i-1}$ have been defined. If $i = 8pd$, then we are done; therefore we may assume that $i < 8pd$ and hence $\mathcal{B}_{i-1}$ is
a \((k, c')\)-box array order-contained in \(B\) with \(c' \geq c/m^{i-1}\). We let

\[
B' = B_1^{i-1}, \ldots, B_{2p}^{i-1}, B_{2p(M(i-1)+1)+1}^{i-1}, \ldots, B_k^{i-1}.
\]

Since \(B_{i-1}\) is \(S\)-smooth and \((S, f_{i-1})\)-optimal, by Lemma 3.2.31, so is \(B'\). Since \((B_{i-1}, f_{i-1})\) is \(p\)-interval, so is \((B', f_{i-1})\) by Lemma 3.2.28.

We now apply Lemma 3.2.33 to \(B'\) with \(f' = f_{i-1}, \mu = 2, \varepsilon, m = 60p\). Note that by construction, \(B'\) is a \((k', c')\)-box array with \(k' \geq k - 2pM(i - 1) \geq 4p + 2pM\) and \(c' \geq 1/m^{i-1}\). It follows that \(f' \leq 1/(60Mp^2)\), and therefore \(f' \leq 1/(2\mu Mp^2)\), and \(f' \leq 1/(60p) \leq \frac{\mu - 1}{6(4p+1)\mu} c\). Furthermore, \(\varepsilon \leq (1/60p)^{8pd+1}c \leq (1/(60p))c' \leq \frac{\mu - 1}{3(4p+1)\mu} c\). Finally, we have \(m = 60p \geq \frac{6(4p+1)\mu}{\mu - 1} c\). This shows that all the conditions of Lemma 3.2.33 are satisfied, and therefore we obtain a set \(A_i\) such that \(A_i \subseteq \bigcup_{j \in \{2p+1, \ldots, 2p(M_1+M_2+1)\}} B'_j\), and a \((k', c'/m)\)-box array \(B''\) order-contained in \(B'\) such that:

- \(B'' = B''_1, \ldots, B''_{k'}\);
- there exists a \(j \in \{1, \ldots, 2p, k' - 2p + 1, \ldots, k'\}\) such that \(A_i\) is either a left \(T_{d,a}(\alpha)\)-cover or a right \(T_{d,b}(\beta)\)-cover of \(B''_j \subseteq B'_j\); and
- for all \(j^* \in \{1, \ldots, k'\} \setminus \{j, 2p+1, 2p+2, \ldots, 2p(M+1)\}\), the set \(A_i\) is anti-complete to \(B'_{j^*}\).

We let \(B_i\) be the box array

\[
B''_1, \ldots, B''_{2p}, B''_{2p+1}, \ldots, B''_{2p(M(i-1)+1)}, B''_{2p+1}, \ldots, B''_{k'}.
\]

It follows that \(B_i\) is a \((k, c'/m)\)-box array order-contained in \(B\) and in \(B_{i-1}\). We let \(f_i = f_{i-1} \cdot m\). By Lemma 3.2.28, since \(f_i \leq \sqrt{f}\) and \(f_i c/m^i \geq cf\), and since \((B, f)\) is \(p\)-interval, it follows that \((B_i, f_i)\) is \(p\)-interval as well. By Lemma 3.2.31 and since \(f_i c/m^i \geq cf\), it follows that \(B_i\) is \(S\)-smooth and \((S, f_i)\)-optimal.
By Lemma 3.2.33, there is a \( j \in \{1, \ldots, 2p, k' - 2p + 1, \ldots, k'\} \) be such that \( A_i \) is either a left \( T_{d,a}(\alpha) \)-cover or a right \( T_{d,b}(\beta) \)-cover of \( B_j'' \subseteq B_j' \). If \( j \leq 2p \), then \( A_i \) is a left \( T_{d,a}(\alpha) \)-cover or a right \( T_{d,b}(\beta) \)-cover of \( B_j'' \). If \( j > 2p \), then \( j \) is a left \( T_{d,a}(\alpha) \)-cover or a right \( T_{d,b}(\beta) \)-cover of \( B_j'' \subseteq B_j + 2(i-1)M \), and \( j + 2(i-1)M \geq k' - 2p + 2(i-1)M = k - 2p + 1 \). It follows that there is a \( j' \in R \) such that \( A_i \) is a left \( T_{d,a}(\alpha) \)-cover or a right \( T_{d,b}(\beta) \)-cover of \( B_j'' \). Since \( \{k, c/m^i\} \)-box array \( B_i \) is order-contained in the \( \{k, c/m^j\} \)-box array \( B_j \) for all \( j \in \{1, \ldots, i - 1\} \), it follows that for all \( j \in \{1, \ldots, i - 1\} \), since \( A_j \) covers \( B_j'' \) for some \( j' \in R \), that \( A_j \) covers \( B_j'' \), and \( A_j \) is anticomplete to \( B_j'' \), for all \( j'' \in \{1, \ldots, 2p, 2p((i-1)M + 1) + 1, \ldots, k\} \setminus \{j'\} \). Since

\[
A_i \subseteq \bigcup_{j \in \{2p+1, \ldots, 2p(M+1)\}} B_j',
\]

\[
\subseteq B_{2p(i-1)M+1} \cup \cdots \cup B_{2p(iM+1)}
\]

it follows that \( A_j \) is anticomplete to \( A_i \) for all \( j \in \{1, \ldots, i - 1\} \). This proves that \( B_i, A_i, f_i \) satisfy all the conditions of (3.5). Now (3.5) follows by repeating this procedure.

Now let \( B' = B_{8pd} = B_1', \ldots, B_k' \) and \( A_1, \ldots, A_{8pd} \) as in (3.5). Note that by construction, it follows that for all \( j \in \{1, \ldots, k\} \setminus R \), there is at most one \( i \in \{1, \ldots, 8pd\} \) with \( A_i \cap B_j \neq \emptyset \), and furthermore, \( A_i \cap B_j = \emptyset \) for all \( i \in \{1, \ldots, 8pd\} \) and \( j \in R \).

For \( j \in R \), we let \( C(j) = \{i \in \{1, \ldots, 8pd\} : j'(8pd, i) = j\} \). Since \( |R| = 4p \), it follows that there is a \( j \in R \) such that \( |C(j)| \geq 2d \). We consider two cases.

Suppose first that for at least \( d \) distinct \( i \in C(j) \), \( A_i \) right \( T_{d,b}(\beta) \)-covers \( B_j' \); let \( C' \subseteq C(j) \) be the set of such \( i \). Since \( c/m^i_{8pd} \geq \varepsilon \), it follows that \( B_j' \neq \emptyset \). Let \( v \in B_j' \),
and for every $i \in C'$, let $T(i)$ be a right-rooted $T_{d,b}(\beta)$ in $G|A_i$ such that $v$ is adjacent to the root $v_i$ of $T(i)$, and non-adjacent to every vertex in $V(T(i)) \setminus \{v\}$.

Since $A_i$ is anticomplete to $A_{i'}$ for all distinct $i, i' \in C'$, it follows that $H = G|\{(v) \cup \bigcup_{i \in C'} V(T(i))\}$ contains $T_{d+b+1}$ as an induced subgraph (since $v$ has at least $d$ neighbors, each of which is the root of a distinct $T(i)$ with $i \in C'$). Furthermore, each $T(i)$ is $B$-rainbow, and for every $j' \in \{1, \ldots, k\} \setminus R$, there is at most one $i \in C'$ such that $V(T(i)) \cap B_{j'} \neq \emptyset$, and $V(T(i)) \cap B_{j'} = \emptyset$ for all $i \in C'$ and $j' \in R$. This implies that $H$ is $B$-rainbow. If $j \in \{1, \ldots, 2p\}$, then $H$ contains a left-rooted $T_{d,b+1}$; otherwise $H$ contains a right-rooted $T_{d,b+1}$. In both cases, the statement of the lemma follows.

Therefore, we may assume that for at least $d+1$ distinct $i \in C(j)$, $A_i$ right $T_{d,a}(\alpha)$-covers $B'_j$. Let $C' \subseteq C(j)$ be the set of such $i$. Since $c/m^{8pd} \geq \varepsilon$, it follows that $B'_j \neq \emptyset$. Let $v \in B'_j$, and for every $i \in C'$, let $T(i)$ be a left-rooted $T_{d,a}(\alpha)$ in $G|A_i$ such that $v$ is adjacent to the root $v_i$ of $T(i)$, and non-adjacent to every vertex in $V(T(i)) \setminus \{v\}$.

Since $A_i$ is anticomplete to $A_{i'}$ for all distinct $i, i' \in C'$, it follows that $H = G|\{(v) \cup \bigcup_{i \in C'} V(T(i))\}$ contains $T_{d+1,a+1}$ as an induced subgraph (since $v$ has at least $d+1$ neighbors, each of which is the root of a distinct $T(i)$ with $i \in C'$). Furthermore, each $T(i)$ is $B$-rainbow, and for every $j' \in \{1, \ldots, k\} \setminus R$, there is at most one $i \in C'$ such that $V(T(i)) \cap B_{j'} \neq \emptyset$, and $V(T(i)) \cap B_{j'} = \emptyset$ for all $i \in C'$ and $j' \in R$. This implies that $H$ is $B$-rainbow. If $j \in \{1, \ldots, 2p\}$, then $H$ contains a left-rooted $T_{d+1,a+1}$; otherwise $H$ contains a right-rooted $T_{d+1,a+1}$. In the former case, the statement of the lemma follows. Therefore, we may assume that the latter case holds.

Let $i^*$ be the smallest element in $C'$, and let $v^*$ be the root of $T(i^*)$. Let $j^* \in \{1, \ldots, k\}$ such that $v^* \in B_{j^*}$. It follows that for all $w \in V(H) \setminus \{v^*\}$, there is a $j' \in \{1, \ldots, k\}$ with $w \in B_j$ and $j > j^*$. Let $H' = H \setminus V(T(i^*))$, then $H'$ contains an
induced $T_{d,a}$ with root $v$. But now $T(i^*)$ is a $T_{d,a}(\alpha)$ with root $v^*$, and therefore, the graph $H^* = G|(V(H') \cup V(T(i^*)))$ contains $T_{d,a}(\alpha + 1)$ (by adding the $T_{d,a}$ with root $v$ as a big branch). Since this is a left-rooted $T_{d,a}(\alpha + 1)$, and since $B$ is $S$-smooth, it follows that $T_{d,a}(\alpha + 1)$ is left-good; this concludes the proof.

The following lemma is an easy application of Lemma 3.2.34.

**Lemma 3.2.35.** Let $T$ be a tree, $p \in \mathbb{N}$. Let $d, h \in \mathbb{N}$. Let $S$ denote the set of all trees $T_{d,h'}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d - 1\}$. Let $k \in \mathbb{N}$ and $f, c > 0$ with $f \leq 1/2$. Let $\varepsilon \leq c/4$ and let $G$ be $\varepsilon$-coherent.

Let $B$ be a $(k,c)$-box array in $G$ such that $B$ is $S$-smooth, $(S,f)$-optimal, and $(B,f)$ is 1-optimal and $p$-interval. Let $M = 2|V(T_{d,h})|$. Suppose that

- $k \geq 16p^2dM + 4p$;
- $f \leq (1/(60Mp^2) \cdot (60p)^{-8pd})^2$; and
- $\varepsilon \leq 1/(60p)^{8pd+1}c$.

Then $G$ contains a $B$-rainbow induced subgraph isomorphic to $T$.

**Proof.** Will will prove that $B$ contains either a left-rooted or a right-rooted $T_{d,h}$. Let $a, i$ be chosen as follows: $a \in \{0, \ldots, h\}$ is maximum subject to $T_{d,a}$ being left-good for $B$; and $\alpha \in \{0, \ldots, d - 1\}$ is maximum subject to $T_{d,a}(\alpha)$ being left-good for $B$. Let $b, j$ be chosen as follows: $b \in \{0, \ldots, h\}$ is maximum subject to $T_{d,b}$ being right-good for $B$; and $\beta \in \{0, \ldots, d - 1\}$ is maximum subject to $T_{d,b}(\beta)$ being right-good for $B$. Clearly, $T_{d,0}$ is left-good and right-good, and so $a, b, \alpha, \beta$ are well-defined.

We may assume the following:

- $a, b < h$ (for otherwise $T_{d,a}$ or $T_{d,b}$ contains $T$ and the statement of the lemma follows);
• $T_{d,a+1}, T_{d,b+1}, T_{d,a}(\alpha+1), T_{d,b}(\beta+1) \in S$ (since either $i+1 \leq d-1$ or $T_{d,a}(\alpha+1)$ is isomorphic to $T_{d,a+1}$);

• $T_{d,a+1}, T_{d,a}(\alpha+1)$ are not left-good; and

• $T_{d,b+1}, T_{d,b}(\beta+1)$ are not right-good.

By symmetry, we may further assume that $a \leq b$.

We now apply Lemma 3.2.34. Since $|V(T_{d,a}(\alpha))| + |V(T_{d,b}(\beta))| \leq M$, it follows that either $T_{d,a+1}$ or $T_{d,b+1}$ or $T_{d,a}(\alpha+1)$ is left-good, or $T_{d,b+1}$ is right-good. Three of these outcomes contradict the above assumptions; the fourth, that $T_{d,b+1}$ is left-good, implies that $T_{d,a+1}$ is left-good (since $T_{d,b+1}$ contains $T_{d,a+1}$). This is a contradiction, and the statement of the lemma follows.

We combine all the results of this section to obtain the following result.

**Lemma 3.2.36.** Let $T$ be a tree; let $k \in \mathbb{N}, k \geq 2$. Then there exist $K, C, C' \in \mathbb{N}$ such that for every graph $G$ and for all $c > 0$, if $\varepsilon \leq c/C'$ and if $G$ is $\varepsilon$-coherent, and $G$ has a $(K, C \cdot c)$-box array $B$, then either $B$ has a $(k, c)$-rainbow matching, or $G$ contains a $B$-rainbow induced subgraph isomorphic to $T$.

**Proof.** Let $d, h \in \mathbb{N}$ be such that $T$ is an induced subgraph of $T_{d,h}$. Let $M = 2|V(T_{d,h})|$ and let $\mathcal{S}$ denote the set of all trees $T_{d,h'}(i)$ for $h' \in \{0, \ldots, h\}$ and $i \in \{0, \ldots, d-1\}$. Let $K_1 = 16k^2dM + 4k$. Let $K_2 = K(K_1, d, h)$, where $K(K_1, d, h)$ is the function from Lemma 3.2.30. Let $K_3 = k(K_2, 2k+2)$, where $k(K_2, 2k+2)$ is the function from Lemma 3.2.26.

Let $X = 2(h+1)dk^{M/2}$. Let $f_3 = (1/(60k^2) \cdot (60k)^{-8kd})^2$ and $f_2 = f_3^{2X}$ and $f_1 = f_2^{2(f2X+2)K_3K_3^{-1}}$. Let $C_2 = 1/f_3X = 1/\sqrt{T_2}$ and $C_1 = C_2f_2/f_1$. Let $C' = (1/(60k))^{8kd+1}$. We claim that choosing this $C'$, along with $K = K_3$ and $C = C_2$ works.

Let $B$ be a $(K_3, C_2 \cdot c)$-box array. By Lemma 3.2.24 applied to $B, k, K_3, f_1$ and $C_1 \cdot c$, it follows there exists a $(K_3, c')$-box array $B'$ order-contained in $B$ such that $(B', f_2)$ is $(2X+2)$-optimal, and $c' \geq cC_1f_1/f_2 \geq cC_2$.
By Lemma 3.2.26, \( \mathcal{B} \) order-contains a \((K_2, cC_2)\)-box array \( \mathcal{B}' \) such that \( \mathcal{B}' \) is \((2k + 2)\)-homogeneous and \((2X + 2)\)-optimal.

By Lemma 3.2.25, if \( T(\mathcal{B}', f_2) \) has a pseudo-matching of size \( k \), then there is a \((k, c'')\)-rainbow matching in \( \mathcal{B} \) with \( c'' = cC_2\sqrt{f_2} \geq c \), and we are done; therefore we may assume that \( T(\mathcal{B}', f_2) \) has no pseudo-matching of size \( k \).

By Lemma 3.2.27, since \( T(\mathcal{B}', f_2) \) has no pseudo-matching of size \( k \), it follows that \( (\mathcal{B}', f_2) \) is \( k \)-interval.

By Lemma 3.2.29, there is a \((K_2, c^*)\)-box array \( \mathcal{B}^* \) such that \( \mathcal{B}^* \) is order-contained in \( \mathcal{B}' \), and \( \mathcal{B}^* \) is \((S, f_3)\)-optimal and \( c^* \geq cC_2f_3^{2(h+1)dk^{3/2}} \geq c \).

Since \( c\sqrt{f_2} \geq cC_2\sqrt{f_2} = cC_2f_2 \), and since \( (\mathcal{B}', f_2) \) is \((2X + 2)\)-optimal and \( \mathcal{B}' \) is a \((K_2, cC_2)\)-box array, it follows that \( (\mathcal{B}', \sqrt{f_2}) \) is \((2X + 1)\)-optimal and \( k \)-interval by Lemma 3.2.28.

By Lemma 3.2.30, it follows that \( \mathcal{B}^* \) order-contains a \((K_1, c)\)-box array \( \mathcal{B}^\sharp \) such that \( \mathcal{B}^\sharp \) is \( S \)-smooth. By Lemma 3.2.28, it follows that \( (\mathcal{B}^\sharp, \sqrt{f_2}) \) is \((2X + 1)\)-optimal and \( k \)-interval. Since \( f_3 \leq \sqrt{f_2}^{2X} \), and since \( f_3c \geq \sqrt{f_2}C_2c \), it follows that \( (\mathcal{B}^\sharp, f_3) \) is \( 1 \)-optimal and \( k \)-interval.

By Lemma 3.2.31, it follows that \( \mathcal{B}^\sharp \) is \((S, f_3)\)-optimal. Since \( G \) is \( \varepsilon \)-coherent, and \( \varepsilon \leq c/C' \leq (1/(60k))^{8kd+1}c \), it follows that \( T, k, d, h, K_1, f_3, c, \varepsilon, G, \mathcal{B}^\sharp \) satisfy the conditions of Lemma 3.2.35, and so \( G \) contains a \( \mathcal{B} \)-rainbow induced subgraph isomorphic to \( T \).

We are now ready to prove the main result of this section, which shows that Conjecture 3.2.3 is true.

**Theorem 3.2.37.** Let \( T \) be a tree. Then there exists an \( \varepsilon > 0 \) such that every \( \varepsilon \)-coherent graph \( G \) contains \( T \) as an induced subgraph.
Proof. We prove the following claim first.

For every tree $T$, there exist $k, c_1, c_2$ such that for every $c > 0$, if $G$ is

\begin{equation}
\varepsilon\text{-coherent for some } \varepsilon \in (0, c/c_2], \text{ and } B \text{ is a } (k, c_1c)\text{-box array in } G, \text{ then}
\end{equation}

$G$ contains a $B$-rainbow induced subgraph isomorphic to $T$.

We prove this by induction on $|V(T)|$. Clearly, (3.6) holds with $k = c = c' = 1$ in
the case that $|V(T)| = 1$. Now suppose that $|V(T)| > 1$, and let $l \in V(T)$ be a leaf
of $T$.

We may assume that (3.6) holds for $T \setminus l$ with constants $k, c_1, c_2$. Now let $K, C, C'$
as in Lemma 3.2.36 applied to $k$ and $T$. Let $k' = K, c'_1 = c_1 C$, and $c'_2 = c_2 C'$. Let
$c > 0$, let $G$ be a graph, $\varepsilon > 0$ such that $G$ is $\varepsilon$-coherent and $\varepsilon \leq c/c'_2$, and let $B$
be a $(k', c'_1c)$-box array. Then, by Lemma 3.2.36, it follows that either $B$ contains
a $B$-rainbow induced subgraph isomorphic to $T$, or $G$ contains a $(k, c_1c)$-rainbow
matching. In the former case we are done; therefore, we may assume that the latter

Let $T = B'_1, \ldots, B'_k$ be a $(k, c_1c)$-box array, and let $A_1, \ldots, A_k$ be cover sets such
that $B'_1, \ldots, B'_k, A_1, \ldots, A_k$ form a $(k, c_1c)$-rainbow matching in $B$. Since $\varepsilon \leq c/c'_2$, it
follows by (3.6) for $T \setminus l$ that $B'$ contains a $B'$-rainbow induced subgraph $H$ isomorphic
to $T \setminus l$. Let $i \in \{1, \ldots, k\}$ be such that vertex $v$ of $H$ corresponding to the neighbor $u$
of $l$ in $T$ is contained in $B'_i$, and let $w$ be a neighbor of $v$ in $A_i$. Then $G|(V(H) \cup \{w\})$
is isomorphic to $T$, since $w$ is adjacent to $v$ and anticomplete to $V(H) \setminus \{v\}$. Moreover,
by the definition of a rainbow matching, it follows that if $j \in \{1, \ldots, k'\}$ is such that
$A_i \subseteq B_j$, then $B'_j \cap B_j = \emptyset$ for all $j' \in \{1, \ldots, k\}$. This implies that $G|(V(H) \cup \{w\})$
is $B$-rainbow, and thus (3.6) is proved.

Now let $T$ be a tree, and let $k, c_1, c_2$ be as in (3.6). Let $c = 1/(2c_1k)$, and let
$\varepsilon = c/c_2$. Let $G$ be $\varepsilon$-coherent. Since $|V(G)| \geq 2c_1k$, we can partition the vertices of
$G$ into $k$ sets $B_1, \ldots, B_k$ with $|B_i| \geq \lfloor |V(G)|/k \rfloor \geq \lfloor |V(G)|/k \rfloor - 1 \geq |V(G)|/k - 1 \geq
Therefore, $B = B_1, \ldots, B_k$ is a $(k, cc_1)$-box array, $\varepsilon \leq c/c_2$, and $G$ is $\varepsilon$-coherent. But then $G$ contains $T$ as an induced subgraph by (3.6). This implies the result of the theorem.

### 3.2.6 Directed graphs

Let $G$ be a digraph; then $G$ contains $H$ as an induced subdigraph if $H$ can be obtained from $G$ by vertex deletion, and $G$ is $H$-free otherwise. A digraph is $\varepsilon$-sparse if its underlying undirected graph is $\varepsilon$-sparse. An (anticomplete) $(x, y)$-pair in $G$ is an (anticomplete) $(x, y)$-pair in the underlying undirected graph of $G$. We use the following notation for small digraphs:

- $\rightarrow\rightarrow$ denotes the digraph with vertex set $\{a, b, c\}$ and edge set $\{(a, b), (b, c)\}$;
- $\rightarrow\leftarrow$ denotes the digraph with vertex set $\{a, b, c\}$ and edge set $\{(a, b), (c, b)\}$; and
- $\leftarrow\rightarrow$ denotes the digraph with vertex set $\{a, b, c\}$ and edge set $\{(b, a), (b, c)\}$.

We consider the following question:

**Question 3.2.38.** For which digraphs $H$ is there an $\varepsilon > 0$ such that every $H$-free $\varepsilon$-sparse digraph $G$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)| > 0$?

By Theorem 3.2.39, it is a necessary condition that the underlying undirected graph $H'$ of $H$ is a forest. But we can say more:

**Theorem 3.2.39.** Let $H = \rightarrow\rightarrow$ be a two-edge directed path. Then, for every $\varepsilon > 0$, and for every sufficiently large $n$, there is an $n$-vertex $H$-free $\varepsilon$-sparse digraph $D$ with no anticomplete $(\varepsilon n, \varepsilon n)$-pair.

To prove this, we need the following definitions. A partially ordered set $P = (S, \prec)$ is a set $S$ and a partial ordering $\prec$ such that:
• for every $x \in S$, $x \neq x$;

• for every $x, y \in S$, if $x < y$ then $y \neq x$; and

• for every $x, y, z \in S$, if $x < y$ and $y < z$, then $x < z$.

If $x < y$ or $y < x$, we say that $y$ is comparable with $x$ (in $P$); otherwise $x$ and $y$ are incomparable (in $P$).

We call $<$ a partial ordering of $S$ if $(S, <)$ is a partially ordered set. A partial ordering $<$ of a set $S$ is linear if for every $x, y \in S$ with $x \neq y$, either $x < y$ or $y < x$.

Two partial orderings $<_1, <_2$ of a set $S$ are compatible if there do not exist $x, y \in S$ such that $x <_1 y$ and $y <_2 x$. A linear extension of a partially ordered set $P = (S, <)$ is a partial ordering $<'$ of $S$ such that $<$ and $<'$ are compatible and $<'$ is linear.

A comparability graph is a graph $G$ such that there is a partially ordered set $P = (V(G), <)$ and $x, y \in V(G)$ are adjacent in $G$ if and only if $x$ and $y$ are comparable in $P$. A directed comparability graph is a digraph $D$ such that there is a partially ordered set $P = (V(D), <)$ and $(x, y) \in E(D)$ if and only if $x < y$. In both cases, $P$ is the underlying partially ordered set of the (directed) comparability graph. It is well-known (see [5, Theorem 6.1.1]) that a digraph is a directed comparability graph if and only if it does not contain $\rightarrow\rightarrow$ as an induced subdigraph.

We use the following result of Fox [25]:

**Theorem 3.2.40** (Fox [25]). Let $\delta \in (0, 1)$. For every sufficiently large positive integer $n$, there is a partially ordered set $P = (S, <)$ on $n$ elements such that no element of $S$ is comparable with $n^\delta$ other elements of $S$, and for every two subsets $A$ and $B$ of $S$ with $|A| = |B| \geq \frac{14n}{\delta \log_2 n}$, there is an element of $A$ that is comparable with an element of $B$.

**Proof of Theorem 3.2.39** Let $\varepsilon > 0$, and $n \in \mathbb{N}$ with $n \geq 2^{20/\varepsilon}/\varepsilon^2$. Let $\delta = 1/2$, and let $P = (S, <)$ be as in Theorem 3.2.40. Let $D$ be the directed comparability graph.
with underlying partially ordered set \( P \). It follows that for every \( v \in S \), \( d(v) \) is the number of \( u \in S \) such that \( v \) is comparable to \( u \). Therefore, \( d(v) \leq n^\delta = \sqrt{n} \leq \varepsilon n \) since \( n \geq 1/\varepsilon^2 \). It remains to show that \( D \) has no anticomplete \((\varepsilon n, \varepsilon n)\)-pair. Suppose that \( A, B \) is an anticomplete pair. By Theorem 3.2.40, it follows that \( \min\{|A|,|B|\} \leq \frac{14n}{\delta \log_2 n} < \varepsilon n \). This concludes the proof.

On the other hand, Theorem 3.2.18 shows that there is a \( \varepsilon > 0 \) such that every \( \varepsilon \)-sparse graph \( G \) with no anticomplete \((\varepsilon n, \varepsilon n)\)-pair contains an induced subdivision of \( K_{2,3} \); but every directed graph with underlying undirected graph isomorphic to a subdivision of \( K_{2,3} \) contains \( \rightarrow \leftarrow \) and \( \leftarrow \rightarrow \) as induced subdigraphs; hence for both of these graphs, Question 3.2.38 can be answered affirmatively. In fact, \((\rightarrow \leftarrow)\)-free graphs have a nice structure:

**Lemma 3.2.41.** Let \( D \) be a \((\rightarrow \leftarrow)\)-free orientation of a graph \( G \). Then there is a sequence of directed graphs \( \emptyset = G_0, G_1, \ldots, G_k = D \) such that \( G_{i+1} \) arises from \( G_i \) by adding a vertex \( v \) such that \( N^+_D(v) = A, N^-_D(v) = B \), and \( A, B \) are cliques, and \( N^+_D(B) \subseteq A \).

**Proof.** If \( G \) is a complete graph, this is trivially true. Let \( v \in V(G) \) such that there exists \( u \in V(G) \) with \( uv \not\in E(G) \). Let \( X \subseteq V(G) \) be maximal such that \( G|X \) is connected and \( V(G) \setminus (X \cup N(X)) \neq \emptyset \) and \( v \in X \). This exists, since \( G \{v\} \) is connected and \( V(G) \setminus N[v] \neq \emptyset \). Let \( Y = N(X), Z = V(G) \setminus (X \cup Y) \).

Suppose that there exist \( b \in Y, c \in Z \) non-adjacent. Then \( X \cup \{b\} \) contradicts the maximality of \( X \) (since \( c \not\in N(X \cup \{b\}) \)). It follows that \( Y \) is complete to \( Z \). Since \( X \neq \emptyset \), it follows that \( |Z| < |V(G)| \). By induction, we may assume that there exists \( c \in Z \) such that \( N^+_D(c), N^-_D(c) \) are cliques.

Let \( A' = N^+(c) \cap Y, B' = N^-(c) \cap Y \). It follows that \( A' \) is a clique. Since \( c \) is anticomplete to \( X \), it follows that for every \( a \in X, b \in B' \) such that \( a \) and \( b \) are adjacent in \( G \), we have that \((b, a) \in E(D)\), since \((c, b) \in E(D)\). Suppose that
b, b′ ∈ B′ are non-adjacent. Let P be a shortest b-b′-path with interior in X. Then there exist a, a′ such that the first edge of P is (b, a) and the last edge of P is (b′, a′). This is impossible in a (→←)-free graph. This proves that B′ is a clique.

It follows that A = A′ ∪ N_{D|Z}^+(c) and B = B′ ∪ N_{D|Z}^−(c) are cliques, and N(c) = A ∪ B. Moreover, suppose that there exists a ∈ V(G) \ (A ∪ B) such that (a, b) ∈ E(G) for some b ∈ B. Then (a, b) and (c, b) form a →←, a contradiction. Thus G arises from G \ v as in the statement of the lemma, and the result follows by induction applied to G \ v.

It follows that every (→←)-free graph has a bisimplicial vertex, i.e. a vertex whose neighborhood is the union of two cliques.

Lemma 3.2.42. Let H be a digraph such that every connected component of H either has size at most two, or is isomorphic to →← or ←→.

Then there exists an ε > 0 such that every ε-sparse H-free digraph G with |V(G)| = n > 1 vertices contains an anticomplete (εn, εn)-pair.

Proof. Let c be the number of components of H, and let H′ be the graph obtained by taking c disjoint copies of K_{2,3}. By Theorem 3.2.18, there exists an ε > 0 such that every ε-sparse graph G with |V(G)| > 1 and no anticomplete (εn, εn)-pair contains an induced subdivision of H′. Now it suffices to prove that every digraph with underlying undirected graph a subdivision of H′ contains H as an induced subdigraph. Consequently, it is enough to show that for every component C of H, every digraph with underlying undirected graph a subdivision of K_{2,3} contains C as an induced subdigraph. This is true if |V(C)| ≤ 2; thus we may assume that C = →← by symmetry. D be a digraph with underlying undirected graph a subdivision of K_{2,3}, and let u, v be the vertices of degree three in D. If u or v has in-degree at least two, then D contains C; thus we may assume that u and v have out-degree at least two.

Now the underlying undirected graph of D is the union of three u-v-paths, and thus
contains a u-v-path \( P = u-w_1-\ldots-w_k-v \) with \((u, w_1) \in E(D)\) and \((v, w_k) \in E(D)\). Since there are \( k + 1 \) edges, each of which contributes 1 to the in-degree of a vertex \( w_i \), but only \( k \) vertices \( w_i \), it follows that there is an \( i \in \{1, \ldots, k\} \) such that \( w_i \) has in-degree two. But then \( D \) contains \( C \). This concludes the proof.

**Theorem 3.2.43.** Let \( C \) be a directed graph with underlying undirected graph \( K_{1,3} \), and such that every edge of \( C \) is oriented away from its leaves. Then, for every \( \varepsilon > 0 \), and for every sufficiently large \( n \), there is an \( n \)-vertex \( C \)-free \( \varepsilon \)-sparse digraph \( D \) with no anticomplete \((\varepsilon n, \varepsilon n)\)-pair.

**Proof.** By Theorem 3.2.5, we can choose a random graph \( G \) with no anticomplete \((\varepsilon n, \varepsilon n)\)-pair and with maximum degree bounded by \( t = 64\frac{1+2\ln 3}{\varepsilon^2} \). Let \( k = 2\lceil 1/\varepsilon \rceil \). We partition the vertex set of \( G \) into \( k \) parts \( V_1, \ldots, V_k \), each of size at most \( n/k + 1 \).

For \( u, v \in V(G) \), we say that \( v \) is to the right of \( u \) if \( u \in V_i \), \( v \in V_j \), and \( i < j \). For \( u \in V(G) \), we let \( R_G(u) = \{ v \in N_G(v) : v \) is to the right of \( u \} \) denote the neighbors to the right of \( u \).

We now define a series of graphs. Let \( H_0 = G \). For \( i \in \{1, \ldots, k\} \), we let \( H_i \) denote the graph obtained from \( H_{i-1} \) by adding all edges \( xy \) such that there is a vertex \( u \in V_i \) and \( x, y \in R_{H_i}(u) \). We let \( H \) denote the graph obtained from \( H_k \) by adding all edges \( xy \) such that there is an \( i \in \{1, \ldots, k\} \) for which \( x, y \in V_i \). Finally, we define a digraph \( D \) with underlying undirected graph \( H \) by choosing an ordering \( \prec \) of the vertices of \( G \), and by adding:

- an edge \((u, v)\) if \( v \in R_H(u) \); and
- an edge \((u, v)\) if there is an \( i \in \{1, \ldots, k\} \) such that \( u, v \in V_i \), and \( u \prec v \).

We claim that \( D \) satisfies the conditions of the theorem if \( n \) is sufficiently large. Since \( G \) has no anticomplete \((\varepsilon n, \varepsilon n)\)-pair, it follows that \( D \) has no anticomplete \((\varepsilon n, \varepsilon n)\)-pair. We claim that \( \Delta(H_i) \leq \Delta(H_{i-1})^2 \) for all \( i \in \{1, \ldots, k\} \). To prove
this, we let $x \in V(H_i)$. Then, for every $y \in N_{H_i}(x) \setminus N_{H_{i-1}}(y)$, there exists a vertex $u \in V(G)$ such that $N_{H_{i-1}}(u) \ni x, y$. This implies that

$$|N_{H_i}(x) \setminus N_{H_{i-1}}(y)| \leq \sum_{u : x \in R_{H_{i-1}}(u)} (|R_{H_{i-1}}(u)| - 1) \leq \Delta(H_{i-1}) (\Delta(H_{i-1}) - 1),$$

and the claim follows. In particular, this implies that $\Delta(H_k) \leq \Delta(G)^{2^k} = t^{2^k}$. Consequently, $\Delta(H) + 1 \leq 1 + t^{2^k} + n/k \leq 1 + t^{2^k} + \varepsilon n/2 \leq \varepsilon n$ for $n$ sufficiently large, and thus $D$ is $\varepsilon$-sparse.

Finally, suppose that $D$ contains $C$ as an induced subdigraph. Let $v$ be the vertex of $D$ that has degree three in $C$. Let $i \in \{1, \ldots, k\}$ be such that $v \in V_i$. By construction, it follows that the remaining vertices of $C$ correspond to vertices in $V_i$ or to the right of $v$. Since $V_i$ is a clique in $H$, it follows that at least two of the vertices in $C$ are to the right of $v$; say $x$ and $y$. Since $x$ and $y$ are non-adjacent, it follows that either $vx$ or $vy$ is not in $E(H_{i-1})$, say $vx \notin E(H_{i-1})$. Since every edge in $E(H_k) \setminus E(H_{i-1})$ has both ends in $V_{i+1} \cup \cdots \cup V_k$ by construction, it follows that $vx \notin E(H_{i-1})$. But $x$ is to the right of $v$, and so $x \notin V_i$. It follows that $vx \notin E(H)$, and so $(v, x) \notin E(D)$, a contradiction. This proves that $D$ is $C$-free.

A directed graph $H$ with underlying undirected graph $P_k$ is an alternating path if $H$ is $(-\rightarrow\rightarrow)$-free. We have resolved Question 3.2.38 for all digraphs $H$ except forests whose components are alternating paths, and with at least one component of size at least three.

### 3.2.7 Holes of fixed length mod $k$

The main result of this section is the following, an application of Theorem 3.2.18.
Theorem 3.2.44. Let $j, k \in \mathbb{N}$ with $j < k$ and such that either $j$ is even or $k$ is odd. Let $C_k^j$ be the set of all holes $C_l$ such that $l$ is congruent to $j$ modulo $k$. Then there exists an $\varepsilon > 0$ such that no $C_k^j$-free graph is $\varepsilon$-coherent.

Proof. We let $t \in \mathbb{N}$ such that $2t \equiv j \mod k$ (choosing $t = j/2$ if $j$ is even, and $t = (k + j)/2$ if $k$ and $j$ are odd). We let $r = R(k, \ldots, k)$ be the Ramsey number of $c_1, \ldots, c_k$, where $c_1 = \cdots = c_k = k$, as defined in Theorem 1.3.1. Next, we let $s = R(r, \ldots, r)$ be the Ramsey number of $d_1, \ldots, d_k$, where $d_1 = \cdots = d_k = r$.

Now let $m = (t + 2)s$, let $H = K_m$, and let $P$ be a Hamilton path in $H$. By Theorem 3.2.18 there exists an $\varepsilon > 0$ such that no graph that does not contain an induced $P$-subdivision of $H$ is $\varepsilon$-coherent. Therefore, it suffices to show that every $P$-subdivision of $H$ contains a graph in $C_k^j$.

Let $H'$ be a $P$-subdivision of $H$. Let $v_1^1, \ldots, v_{t+2}^1, v_1^2, \ldots, v_{t+2}^2, \ldots, v_1^s, \ldots, v_{t+2}^s$ denote the vertices of $P$ in order. For every $a, b \in \{1, \ldots, s\}$ with $a \neq b$, and for $c \in \{1, \ldots, t + 2\}$, we let $P_{c}^{a,b}$ denote the $v_c^a$-$v_c^b$-path in $H'$ corresponding to the edge $v_c^a v_c^b$ in $H$.

We let $J = K_s$, $V(J) = w_1, \ldots, w_s$, and we let $f : E(J) \to \{0, \ldots, k - 1\}$, where $f(w_aw_b)$ is defined as $L \mod k$, where $L$ is the length of $P_{1}^{a,b}$. By Theorem 1.3.1 it follows that there exists a set $I \subseteq \{1, \ldots, s\}$ with $|I| \geq r$, and a $p \in \{0, \ldots, k - 1\}$ such that $f(w_aw_b) = p$ for all $a, b \in I$.

Now let $J' = K_p$, $V(J') = \{u_i : i \in I\}$. We let $f : E(J') \to \{0, \ldots, k - 1\}$, where $f(u_au_b)$ is defined as $L \mod k$, where $L$ is the length of $P_{t+1}^{a,b}$. By Theorem 1.3.1 it follows that there exists a set $I' \subseteq I$ with $|I'| \geq k$, and a $p' \in \{0, \ldots, k - 1\}$ such that $f(u_au_b) = p'$ for all $a, b \in I'$.  

118
We let \( I' = \{i_1, \ldots, i_k\} \). Let

\[
\begin{align*}
P_1 &= v_{i_1}^{i_1} - v_{i_2}^{i_2} - \ldots - v_{i_t}^{i_t} \\
P_2 &= v_{i_k}^{i_k} - v_{i_2}^{i_2} - \ldots - v_{i_t}^{i_t} \\
P_3 &= v_{i_1}^{i_1} P_{i_1, i_2}^{i_1, i_2} v_{i_2}^{i_2} P_{i_2, i_3}^{i_2, i_3} \ldots P_{i_k-1, i_k}^{i_k-1, i_k} v_{i_k}^{i_k} \\
P_4 &= v_{i_t+1}^{i_t+1} P_{i_t+1, i_t+1}^{i_t+1, i_t+1} v_{i_t+1}^{i_t+1} P_{i_t+1, i_t+1}^{i_t+1, i_t+1} \ldots P_{i_k-1, i_k}^{i_k-1, i_k} v_{i_k}^{i_k}.
\end{align*}
\]

Then \( P_1, P_2 \) each have length \( t \). The path \( P_3 \) has length congruent to \( kp \equiv 0 \mod k \), and \( P_4 \) has length congruent to \( kp' \equiv 0 \mod k \). It follows that \( Q = v_{i_1}^{i_1} P_1 v_{i_t+1}^{i_t+1} P_4 v_{i_t+1}^{i_t+1} P_2 v_{i_t+1}^{i_t+1} P_3 v_{i_1}^{i_1} \) is a hole of length congruent to \( 2t \mod k \), and therefore \( Q \in C_j^k \). This implies the result.

Theorem 3.2.44 allows us to find holes of all residues \( j \mod k \) for which some hole in \( C_j^k \) is even. This is best possible, as shown by the following result:

**Theorem 3.2.45.** For every \( \varepsilon > 0 \), and for every sufficiently large \( n \), there exists an \( n \)-vertex \( \varepsilon \)-sparse perfect graph with no \((\varepsilon n, \varepsilon n)\)-anticomplete pair.

**Proof.** This follows directly from Theorem 3.2.39, since the underlying undirected graph of a \((\rightarrow\rightarrow)\)-free graph is a comparability graph and hence perfect (see, for example, [36]).

\( \square \)

### 3.3 Linear and polynomial anticomplete pairs in sparse graphs

In this section, we consider the following questions:

**Question 3.3.1.** For which graphs \( H \) does there exist a constant \( \varepsilon > 0 \) such that every \( \varepsilon \)-sparse \( H \)-free graph \( G \) with \( |V(G)| = n \) contains an anticomplete \((\varepsilon n^\varepsilon, \varepsilon n)\)-pair?
**Question 3.3.2.** For which graphs $H$ do there exist constants $k, \varepsilon > 0$ such that for all $c$ with $1/2 \geq c > 0$, every $\varepsilon$-sparse $H$-free graph $G$ with $|V(G)| = n$ contains a $c$-sparse $(\varepsilon c^kn, \varepsilon n)$-pair?

The motivation for Question 3.3.2 is the following result:

**Lemma 3.3.3** (Chudnovsky, Fox, Scott, Seymour, Spirkl [9]). Let $H, H'$ be graphs such that $H$ and $H'$ satisfy Question 3.3.2. Then there exists a $\delta > 0$ such that every $\{H, H'\}$-free graph $G$ has a clique or a stable set of size at least $2e^{\sqrt{\log n \log \log n}}$.

In particular, we will prove that Question 3.3.2 holds for $H = C_5 = C_5^c$, and therefore improve Theorem 3.1.6 for $H = C_5$.

Conlon, Fox, and Sudakov [18] conjectured that for both questions, every graph $H$ should have these properties, and proved that graphs $H$ with the second property also have the first. We include a proof for completeness.

**Lemma 3.3.4** (Conlon, Fox, Sudakov [18]). Let $H$ be a graph such that Question 3.3.2 is true for $H$. Then Question 3.3.1 is true for $H$.

**Proof.** Let $k, \varepsilon$ be as in Question 3.3.2 for $H$. We may assume that $k + 1 \leq 1/(2\varepsilon)$ and $\varepsilon \leq 4^{-(4k+1)}$. We let $c = 4^{1/k}n^{(\varepsilon - 1)/k}$. Let $G$ be an $\varepsilon$-sparse $H$-free graph with $|V(G)| = n$. Then $G$ contains a $c$-sparse $(\varepsilon c^kn, \varepsilon n)$-pair $A, B$. It follows that $|A| \geq \varepsilon c^kn \geq 4\varepsilon n^\varepsilon$. By Lemma 3.2.6 we know that $n \geq 1/\varepsilon \geq 4^{4k+1}$.

Since $A, B$ is $c$-sparse, it follows that for $E = \{e \in E(G) : e = uv, u \in A, v \in B\}$, we have $|E| \leq c|A||B|$. Let $A' \subseteq A$ be the set of vertices $a \in A$ with $|N(a) \cap B| \geq 4c|B|$. Since $c|A||B| \geq |E| \geq 4c|B||A'|$, it follows that $|A'| \leq |A|/4$. If $\lceil \varepsilon n^\varepsilon \rceil = 1$, then $x, V(G) \setminus N[x]$ is an anticomplete $(\varepsilon n^\varepsilon, (1 - \varepsilon)n)$-pair for every $x \in V(G)$, and the result follows. Therefore, we assume that $\varepsilon n^\varepsilon \geq 1$ from now on. Therefore, $|A \setminus A'| \geq 3|A|/4 \geq |A|/4 + 1 \geq \lceil \varepsilon n^\varepsilon \rceil$. Let $A'' \subseteq A \setminus A'$ with $|A''| = \lceil \varepsilon n^\varepsilon \rceil$. 

120
Since \( A'' \cap A' = \emptyset \), it follows that \( N(A'') \cap B \leq 4c|A''| |B| \leq 4c(\varepsilon n^\varepsilon + 1)|B| \geq 8c\varepsilon n^\varepsilon |B| \). But

\[
8c\varepsilon n^\varepsilon = 8\varepsilon 2^{1/k} n^{((k+1)\varepsilon -1)/k} \leq 8\varepsilon 2^{1/k} n^{-1/(2k)} = 8 \cdot (4/n)^{1/(2k)},
\]

and since \( n \geq 4^{4k+1} \), it follows \( 8 \cdot (4/n)^{1/(2k)} \leq 8(4^{-4k})^{1/2k} = 8/16 = 1/2 \). So \( A'', B \setminus N(A'') \) is an anticomplete \((\varepsilon n^\varepsilon, \varepsilon n)\)-pair, and consequently, \( H \) satisfies Question 3.3.1 with \( \varepsilon/2 \) instead of \( \varepsilon \).

We make following simple observation, which allows us to apply results from Section 3.2 in this context:

**Lemma 3.3.5.** Let \( \mathcal{C} \) be a class of graphs such that for all \( \varepsilon > 0 \), there exist \( c \in (0,1] \) and \( G \in \mathcal{C} \) such that \( G \) is \( \varepsilon \)-sparse and has no \( c \)-sparse \((\varepsilon c^k n, \varepsilon n)\)-pair. Then \( G \) is \( \varepsilon \)-coherent.

**Proof.** An anticomplete \((\varepsilon n, \varepsilon n)\)-pair is a \( c \)-sparse \((\varepsilon c^k n, \varepsilon n)\)-pair for all \( c \in (0,1] \); this implies the result.

This implies that the property in Question 3.3.2 is true for \( H \) if there is an \( \varepsilon > 0 \) such that no \( \varepsilon \)-coherent graph is \( H \)-free. Therefore, Question 3.3.2 is true for subdivided caterpillars by Theorem 3.2.23 and for all trees by Theorem 3.2.37. In Section 3.3.1, we prove that Question 3.3.2 holds for every hole \( H = C_k \). This leads to an improvement of Theorem 3.1.6 for \( H = C_5 \). In Section 3.3.3, we show that for every fixed long subdivision \( H \) of a graph \( H' \) (meaning every edge of \( H' \) is subdivided at least a certain number of times), Question 3.3.2 is true for \( H \). In Section 3.5, we prove that Question 3.3.2 is true for all bipartite graphs, and for “almost bipartite” graphs, which are a special class of triangle-free graphs.

One of the simplest graphs for which Question 3.3.2 is still open is the triangle \( K_3 \); here the best known result is given in Section 3.7.
A crucial difference between $\varepsilon$-coherent graphs and $\varepsilon$-sparse graphs with no $c$-sparse $(\varepsilon c^k n, \varepsilon n)$-pair is that the latter are not $(\varepsilon, 2)$-coherent, as shown in the following lemmas.

**Lemma 3.3.6.** Let $1/2 \geq c > 0$, $1/2 > \varepsilon > 0$, and let $G$ be an $\varepsilon$-sparse graph with no $c$-sparse $(\varepsilon c^k n, \varepsilon n)$-pair. Then $|V(G)| > 1/(\varepsilon c^k)$.

*Proof.* Suppose that $|V(G)| \leq 1/(\varepsilon c^k)$. Then $\{x\}, V(G) \setminus N[x]$ is an anticomplete $(\varepsilon c^k n, (1 - \varepsilon)n)$-pair, a contradiction. This proves the lemma. \hfill $\Box$

**Lemma 3.3.7.** Let $1/2 \geq c > 0$, $1/4 > \varepsilon > 0$ and let $G$ be an $\varepsilon$-sparse graph with no $c$-sparse $(\varepsilon c^k n, \varepsilon n)$-pair. Then $G$ has a vertex of degree at least $cn/2$.

*Proof.* Since $\varepsilon \leq 1/2$, it follows that $|V(G)| \geq 2$. Let $A \subseteq V(G)$ with $|A| = \lfloor |V(G)|/2 \rfloor$. Since $|A| \geq |V(G)|/4$, it follows that $A, V(G) \setminus A$ is not a $c$-sparse pair. It follows that $|E(A, V(G) \setminus A)| \geq c|A||V(G) \setminus A|$. Therefore, there is a vertex $v$ in $A$ with at least $c|V(G) \setminus A|$ neighbors in $V(G) \setminus A$, so $d(v) \geq cn/2$. \hfill $\Box$

**Lemma 3.3.8.** Let $G$ be a graph, $A, B \subseteq V(G)$. If $A, B$ is not a $c$-sparse $(|A|, |B|)$-pair, then there is a vertex $v \in A$ with $|N(v) \cap B| \geq c|B|$.

*Proof.* Since $A, B$ is not $c$-sparse, it follows that

$$c|A||B| \leq |E(A, B)| = \sum_{a \in A} |N(a) \cap B| \leq |A| \max_{a \in A} |N(a) \cap B|.$$  

This implies that $\max_{a \in A} |N(a) \cap B| \geq c|B|$, as claimed. \hfill $\Box$

### 3.3.1 Holes

The goal of this section is to prove the following result:

**Theorem 3.3.9.** Let $l \geq 4$. Then there exist $k, \varepsilon > 0$ such that for every $c$ with $1/2 \geq c > 0$, every $\varepsilon$-sparse $C_l$-free graph $G$ with $|V(G)| = n$ contains an anticomplete $(\varepsilon c^k n, \varepsilon n)$-pair. This is true with $k = 1$ when $l \geq 5$. 

122
This implies that Question 3.3.2 is true when $H$ is a hole. We prove Theorem 3.3.9 separately for $l = 4$ in Lemma 3.3.11 and for $l \geq 5$ in Lemma 3.3.10.

**Lemma 3.3.10.** Let $l \geq 5$. Then there exist $k, \varepsilon > 0$ such that for every $c$ with $1/2 \geq c > 0$, every $\varepsilon$-sparse $C_l$-free graph $G$ with $|V(G)| = n$ contains an anticomplete $(\varepsilon cn, \varepsilon n)$-pair.

**Proof.** Let $c > 0$ and $l \geq 5$. We let $k = 1$ and $\varepsilon = \frac{1}{4l-2}$; it follows that $\varepsilon \leq \frac{1}{8}$. We first prove:

If $\varepsilon \leq 1/8$ and $k \geq 1$, then there exist $u, v \in V(G)$ with $uv \in E(G)$, and sets $A \subseteq N(u) \setminus N(v)$ and $B \subseteq N(v) \setminus N(u)$ such that $A$ is anticomplete to $B$, and $|N(A)| \geq n/4$, and $|B| \geq \varepsilon cn.

(3.7) By Lemma 3.3.7, it follows that there is a vertex $v \in V(G)$ with $|N(v)| \geq cn/2$. Since $N(v), V(G) \setminus N^2[v]$ is an anticomplete pair, it follows that $|V(G) \setminus N^2[v]| \leq \varepsilon n$. Consequently, $|N^2(v)| \geq |V(G) \setminus N[v]| - \varepsilon n \geq |V(G)| - 2\varepsilon n \geq 3n/4$, since $\varepsilon \leq 1/8$.

Now let $B = N^2(v)$, and let $A \subseteq N(v)$ minimal with respect to inclusion subject to $|N(A) \cap B| \geq n/2$. By the minimality of $A$, it follows that $|N(A) \cap B| \leq n/2 + \varepsilon n \leq 5n/8$. Consequently, $C = B \setminus N(A)$ satisfies $|C| \geq 3n/4 - 5n/8 \geq n/8 \geq \varepsilon n$.

Since $A$ is anticomplete to $B$, it follows that $|A| \leq \varepsilon cn$, and therefore $|N(v) \setminus A| \geq \frac{cn}{2} - \varepsilon cn \geq \frac{cn}{4}$. By Lemma 3.3.8, it follows that there is a vertex $u \in N(v)$ with $N(u) \cap C \geq \varepsilon cn$. Let $A' = A \setminus N(u)$. If $|N(A')| \geq n/4$, then the claim follows since $v$ is anticomplete to $N(u) \cap C$, and $u$ is anticomplete to $A'$, and $A'$ is anticomplete to $N(u) \cap C$ since $A$ is anticomplete to $C$. Therefore, we may assume that $N(A') \leq n/4$. Since $|N(A) \cap B| \geq n/2$, it follows that $|N(A \setminus A') \cap B| \geq n/4$.

Let $X = N(u) \cap A = A \setminus A'$, and $Y = N(u) \cap C$. Then $X$ is anticomplete to $Y$, and $5n/8 \geq |N(X) \cap B| \geq n/4$. We let $X' \subseteq X$ be minimal with respect to inclusion subject to $|N(X')| \geq n/4$. It follows that $|V(G) \setminus (N[u] \cup N(X'))| \geq l - 1$, and therefore...
By repeated application of (3.7), we may replace $uv$ and obtain sets $B \subseteq X$ satisfying $|Z| \geq \varepsilon cn$. But now $u$ and $y$ are the desired vertices since $u$ is anticomplete to $Z$, and $y \in Y$ is anticomplete to $X' \subseteq X$, and $X'$ is anticomplete to $Z$. This proves (3.7).

Let $u, v \in V(G)$, and let $P$ be a $u$-$v$-path of length at most $l - 5$. Let $A \subseteq N(u) \setminus N(V(P) \setminus \{u\})$ and $B \subseteq N(v) \setminus N(V(P) \setminus \{v\})$ such that $A$ is anticomplete to $B$, and $|N(A)| \geq n/4$, and $|B| \geq \varepsilon c^n$. Then there is a $u$-$y$-path $P'$ with $P' = Pv-y$, and $A' \subseteq N(u) \setminus N(V(P') \setminus \{u\})$ and $B' \subseteq N(y) \setminus N(V(P') \setminus \{y\})$ such that $A'$ is anticomplete to $B'$, and $|N(A')| \geq n/4$, and $|B'| \geq \varepsilon c^n$.

By removing vertices from $A$, we may assume that $A$ is minimal with respect to inclusion subject to $|N(A)| \geq n/4$. It follows that $|V(G) \setminus (N(A) \cup N(V(P))| \geq |V(G)| - (1/4 + (l - 4)\varepsilon)n \geq n - n/4 - \frac{l - 5}{4l - 2}n \geq n/4$. Since $|B| \geq \varepsilon c^n$, it follows that there is a vertex $y \in B$ with $|N(y) \cap (V(G) \setminus (N(A) \cup N(V(P))|) \geq cn/4 \geq c^l \varepsilon n$. We let $B' = N(y) \cap (V(G) \setminus (N(A) \cup N(V(P))))$, and $P' = Pv-y$. It follows that $y$ is anticomplete to $A$, and that $v$ is anticomplete to $B'$, and $A$ is anticomplete to $B'$. This proves (3.8).

By (3.7), there exist $u, v \in V(G)$ with $uv \in E(G)$, and sets $A \subseteq N(u) \setminus N(v)$ and $B \subseteq N(v) \setminus N(u)$ such that $A$ is anticomplete to $B$, and $|N(A)| \geq n/4$, and $|B| \geq \varepsilon c^n$. By repeated application of (3.7), we may replace $uv$ by a $u$-$y$-path $P'$ of length $l - 4$, and obtain sets $A' \subseteq N(u) \setminus N(V(P') \setminus \{u\})$ and $B' \subseteq N(y) \setminus N(V(P') \setminus \{y\})$ such that $A'$ is anticomplete to $B'$, and $|N(A')| \geq n/4$, and $|B'| \geq \varepsilon c^n$. Let $C = N(A') \setminus N[P']$; it follows that $|C| \geq n/4 - (l - 3)\varepsilon n \geq \frac{l - 1/2 - l + 3}{4l - 2}n \geq 2\varepsilon n$. Since $|B'| \geq \varepsilon c^n$, and $|C| \geq \varepsilon n$, it follows that there is a vertex $z \in B'$ with a neighbor $x \in C \cap N(B')$. Let
Lemma 3.3.11. There exist $k, \varepsilon > 0$ such that for every $c$ with $1/2 \geq c > 0$, every $\varepsilon$-sparse $C_4$-free graph $G$ with $|V(G)| = n$ contains an anticomplete $(\varepsilon c^k n, \varepsilon n)$-pair. We can choose $k = 2$ and $\varepsilon = 1/6$.

Proof. Let $c > 0$, and let $G$ be $C_4$-free. Let $k = 2, \varepsilon = 1/6$. By Lemma 3.3.7, there is a vertex $v \in V(G)$ with $|N(v)| \geq cn/2$. Since $|N(v)| \geq \varepsilon c^k n$, and $|N[v]| \leq \varepsilon n$, it follows that $N(v), V(G) \setminus N[v]$ is not $c$-sparse. Therefore, by Lemma 3.3.8 there is a vertex $w \in V(G) \setminus N[v]$ with $|N(w) \cap N(v)| \geq c^2 n/2$. Since $w$ and $v$ are non-adjacent, and since $G$ is $C_4$-free, it follows that $N(w) \cap N(v)$ is a clique.

Now let $B = V(G) \setminus N[v]$. Since $|N(w) \cap N(v)| \geq c^2 n/2$, it follows that $|B \setminus N(N(w) \cap N(v))| < \varepsilon n$, and so $|B \cap N(N(w) \cap N(v))| \geq 4 \varepsilon n$. We let $A \subseteq N(w) \cap N(v)$ be minimum subject to $|N(A) \cap B| \geq 2 \varepsilon n$. It follows that $|N(A) \cap B| < 3 \varepsilon n$ by the minimality of $A$, and so $|B \setminus N(A)| \geq \varepsilon n$. This implies that $|A| < c^2 n/4$, since $A, B \setminus N(A)$ is an anticomplete pair. It follows that $C = (N(w) \cap N(v)) \setminus A$ contains at least $c^2 n/4 = \varepsilon c^2 n$ vertices. Let $D = B \setminus N(A)$. By Lemma 3.3.8 and since $C, D$ is not $c$-sparse, it follows that there is a vertex $x \in C$ with $|N(x) \cap D| \geq c \varepsilon n$.

We let $E = N(x) \cap D$. It follows that $A$ is anticomplete to $E$. We let $F = (D \cap N(E)) \setminus N[x]$. Since $E$ is anticomplete to $D \setminus N(E)$, it follows that $|D \setminus N(E)| \leq \varepsilon n$. Therefore, $|F| \geq |D| - 2 \varepsilon n = \varepsilon n$. Let $y \in F$, and let $z \in A \cap N(y)$. Let $u \in E \cap N(y)$. It follows that $x$ is adjacent to $u$, since $x$ is complete to $E$; non-adjacent to $y$, since $x$ is anticomplete to $F$; and adjacent to $z$, since $A \cup C$ is a clique. Moreover, $u$ is adjacent to $y$, and non-adjacent to $z$, since $A$ is anticomplete $E \subseteq D$. Finally, $y$ is adjacent to $z$. It follows that $G \{x, y, z, u\}$ is a $C_4$. This proves the lemma. □

Note that $C_4$ is known to have the EH-property by Theorem 3.1.3.
3.3.2 All trees

This section contains a simple proof that question 3.3.2 is true for all trees.

A \((c,d,h)\)-tree is a \((T_{d,h},v)\)-cover structure \(\{A_t\}_{t \in V(T_{d,h})}\) where \(v\) is a root of \(T_{d,h}\), such that \(|A_t| = 1\) for all \(t \in V(T) \setminus \{v\}\), and \(|A_v| \geq c\). For a \((c,d,h)\)-tree \(\{A_t\}_{t \in V(T_{d,h})}\) where \(v\) is a root of \(T_{d,h}\), we call \(A_v\) the root of the \((c,d,h)\)-tree and \(\bigcup_{t \in V(T_{d,h}) \setminus \{v\}} A_t\) the crown of the \((c,d,h)\)-tree.

Lemma 3.3.12. Let \(d \in \mathbb{N}\) and let \(\varepsilon \leq 1/d\) and \(k \geq d\). Let \(G\) be an \(\varepsilon\)-sparse graph with no anticomplete \((\varepsilon c^k n, \varepsilon n)\)-pair. Then \(G\) contains a \((c^d n/2, d, 1)\)-tree.

Proof. We prove this by induction on \(d\). By Lemma 3.3.7, it follows that \(G\) contains a vertex \(v\) of degree at least \(c n/2\). This implies the result for \(d = 1\).

Now let \(d > 1\), and let \(\{A_t\}_{t \in V(K_{1,d-1})}\) be a \((c^{d-1} n/2, d - 1, 1)\)-tree. Let \(v\) be the root of \(K_{1,d-1}\). It follows that \(|A_v| \geq c^{d-1} n/2\). Let \(A = \bigcup_{t \in V(T) \setminus \{v\}} A_t\). It follows that \(A_v\) is complete to \(A\), and hence \(|A_v| \leq \varepsilon n\). Moreover, \(|N[A]| \leq (d - 1)\varepsilon n\), and so \(|G \setminus N[A]| \geq \varepsilon n\). Since \(A_v, G \setminus N[A]\) is not a \(c\)-sparse pair, it follows that there is a vertex \(u \in G \setminus N[A]\) with \(|N(u) \cap A_v| \geq c |A_v| \geq c^d n/2\). We add a vertex \(x\) adjacent to \(v\) to \(K_{1,d-1}\) to obtain \(K_{1,d}\), and let \(A'_v = \{u\}, A'_v = A_v \cap N(u),\) and \(A'_t = A_t\) for all \(t \in V(K_{1,d-1}) \setminus \{v\}\). Then \(\{A'_t\}_{t \in V(K_{1,d})}\) is a \((c^d n/2, d, 1)\)-tree. This implies the result.

Theorem 3.3.13. For every \(d,h \in \mathbb{N}\) there exists an \(\varepsilon > 0\) such that for all \(c > 0\), every \(\varepsilon\)-sparse graph \(G\) with no anticomplete \((\varepsilon c^{d+1} n, \varepsilon n)\)-pair contains \(T_{d,h}\) as an induced subgraph.

Proof. Since \(\{v\}, V(G) \setminus N[v]\) is an anticomplete \((1, (1 - \varepsilon)n)\)-pair, we may assume that \(\varepsilon c^{d+1} n > 1\). Using this assumption, we will prove by induction on \(h\) that \(G\) contains a \((\varepsilon c^d n, d, h)\)-tree. For \(h = 1\), this follows from Lemma 3.3.12 with \(\varepsilon_1 = 1/d\). Now suppose that \(h > 1\) and the result is true for \(h - 1\) with \(\varepsilon_{h-1}\).
We let \( C = d|V(T_d,h-1)|+1 \). By Lemma 3.2.14 for \( T = K_{1,d} \) and for \( v \) chosen as the center of \( T \), \( V(T) = \{v_1, \ldots, v_d, v\} \), there exist \( \varepsilon', \varepsilon > 0 \) such that \( \varepsilon' \geq C \cdot \varepsilon' \) and every \( \varepsilon' \)-coherent graph \( G' \) contains a \((T, v)\)-cover structure \( \{A_t\}_{t \in V(T)} \) with \( |A_t| \geq \varepsilon'n \), where \( n = |V(G)| \). Moreover, every \( A_t \) with \( t \in V(T) \setminus \{v\} \) is chosen according to Lemma 3.2.12 applied to a \((k', d')\)-box array with \( d' \geq \varepsilon' \), and \( |N_{G}(A_t)| \geq \varepsilon'n \).

We let \( \varepsilon' \) as in Lemma 3.2.14 and let \( \varepsilon = \varepsilon_h = \varepsilon' \cdot \varepsilon_{h-1} \), and \( k = 1 \). Let \( G \) be an \( \varepsilon \)-sparse graph, \( c > 0 \), and suppose that \( G \) does not have a \( c \)-sparse \((\varepsilon cn, \varepsilon n)\)-pair. Let \( X \subseteq V(G) \) with \( |X| \geq \varepsilon'n \); then \( G|X \) does not have a \( c \)-sparse \((\varepsilon_{h-1}c^d|X|, \varepsilon_{h-1}|X|)\)-pair, and consequently \( X \) contains an \((\varepsilon_{h-1}c^d|X|, d, h - 1)\)-tree.

Therefore, when constructing \( \{A_t\}_{t \in V(T)} \) using Lemma 3.2.14 we may assume that for every \( t \in V(T) \setminus \{v\} \), \( A_t \) is chosen by applying Lemma 3.2.12 to a set \( B_i \) with \( |B_i| \geq \varepsilon'n \geq C\varepsilon' \geq 2\varepsilon'n \) and a \((k', d')\)-box array \( B_1, \ldots, B_{k'} \) with \( d' \geq C\varepsilon' \). Moreover, \( |B_i| \cdot \varepsilon_{h-1}c^{d+1} \geq \varepsilon_{h-1}c^{d+1}n > 1 \). It follows by induction that \( B_i \) contains an \((\varepsilon_{h-1}c^d|X|, d, h - 1)\)-tree \( T_i \). We let \( A \) denote the root of \( T_i \), and let \( B \) denote the crown of \( T_i \). Since \( |A| \geq \varepsilon_{h-1}c^d|X| \geq 2\varepsilon cn \), it follows that \( |V(G) \setminus N[B]| \leq \varepsilon n \). Now let \( B_i' = A \cup B \) and let \( \{a_1, \ldots, a_s\} = B \). Let \( C_0 = (B_1 \cup \cdots \cup B_{k'}) \setminus B_i \). Since \( k' \geq 2 \), it follows that \( |C_0| \geq \varepsilon'n \geq 2\varepsilon n \).

We now define \( b_1, \ldots, b_m, C_0, C_1, \ldots, C_m, A_0, A_1, \ldots, A_m \) as follows. We let \( A_0 = A \), and let \( i = 1 \), and repeat the following:

- since \( |A_{i-1}| \geq |A| - i \geq \varepsilon c^d n \), and since \( |C_{i-1}| \geq |C_0|/2 \geq \varepsilon n \), it follows that \( A_{i-1}, C_{i-1} \) is not a \( c \)-sparse pair; consequently there is a vertex \( b_i \in A_{i-1} \) with \( |N(b_i) \cap A_{i-1}| > c|C_{i-1}| \geq c|C_0|/2 \);

- let \( b_i \) be as above, \( A_i = A_{i-1} \setminus \{b_i\} \) and \( C_i = C_{i-1} \setminus N(b_i) \);

- if \( i \geq \varepsilon c^d n \), let \( m = i \), and stop;

- if \( |C_i| < |C_0|/2 \), let \( m = i \), and stop.
When this procedure terminates, we have that \(|C_0 \setminus C_m| = |C_0 \cap N(\{b_1, \ldots, b_m\})| > |C_0| - mc|C_0|/2\), and by the last bullet, it follows that \(m \leq 1/c\). Since \(1/c \cdot (\varepsilon d^k n)^{-1} = 1/(\varepsilon d^k+1)n < 1\) by our initial assumption, it follows that at termination, the condition of the second-to-last bullet is not satisfied, and so the condition of the last bullet holds. Therefore, \(|N(\{b_1, \ldots, b_m\})| > |C_0|/2\). Now let \(a + i = b_i\) for \(i = \{1, \ldots, m\}\), and let \(B'_i = \{a_1, \ldots, a_{s+m}\}\). Since \(|(B_i \cup \ldots \cup B_k) \setminus (B_i \cup N(B'_i))| < |(B_i \cup \ldots \cup B_k) \setminus B_i|/2\), this procedure satisfies the conditions of Lemma 3.2.12; this implies that we may assume that every \(A_t\) with \(t \in V(T) \setminus \{v\}\) consists of the crown and a subset of size at most \(1/c\) of the root of a \((\varepsilon d^k n, d, h - 1)\)-tree.

Now, by Lemma 3.2.14 we let \(\{A_t\}_{t \in V(T)}\) be a \((T, v)\)-cover structure such that

- For every \(v_i \in V(T) \setminus \{v\}\), \(A_{v_i}\) consists of the crown \(A^i\) and a subset \(B^i\) of size at most \(1/c\) of the root of a \((\varepsilon d^k n, d, h - 1)\)-tree; and

- \(|A_v| \geq C\varepsilon n\).

We let \(A = A_v \setminus (N(A^1 \cup \ldots \cup A^d))\). Since \(|A^i| \leq |V(T_{d,h})|\) for all \(i \in \{1, \ldots, d\}\), it follows that \(|A| \geq |A_v| - d|V(T_{d,h})| \geq \varepsilon n\). It follows that \(A\) is covered by \(B^i\) for all \(i \in \{1, \ldots, n\}\). For every \(a \in A\), we let \(f(a) = (b_1(a), \ldots, b_d(a))\) with \(b_i(a) \in B^i \cap N(a)\). The number of possible sequences \(f(a)\) is bounded by \(|B^1| \cdot \ldots \cdot |B^d| \leq (1/c)^d\). It follows that there exists a set \(B \subseteq A\) of size at least \(|A|/(1/c)^d \geq \varepsilon d^k n\) such that \(f(a) = f(a') = (b_1, \ldots, b_d)\) for all \(a, a' \in B\).

We now define a \((\varepsilon d^k n, d, h)\)-tree \(T\) by letting \(B\) be the root of \(T\), and by letting \(\{b_1\} \cup A^1 \cup \ldots \cup \{b_d\} \cup A^d\) be the crown \(A\) of \(T\). It follows that \(A\) has \(d\) connected components, each of which is isomorphic to \(T_{d,h-1}\) with vertex set \(\{b_1\} \cup A^i\) and root \(b_i\) for some \(i \in \{1, \ldots, d\}\), and that \(\{b_i\}\) is complete to \(B\), and \(A^i\) is anticomplete to \(B\) for all \(i \in \{1, \ldots, d\}\). Therefore, this is a \((\varepsilon d^k n, d, h)\)-tree, as claimed. \(\square\)
3.3.3 Long exact subdivisions

In this section, we prove that question \ref{question:long_subdivisions} holds for graphs \( H \) that are long subdivisions.

A graph \( H \) is a \((\geq k)\)-subdivision of a graph \( H' \) if \( H \) arises from \( H' \) by subdividing edges, and each edge of \( H' \) corresponds to a path in \( H \) of length at least \( k \).

A \((k,c)\)-star cover in a graph \( G \) consists of the following:

- a subdivided star \( H \) with center \( v \) and leaves \( v_1, \ldots, v_k \), such that the unique \( v-v_i \)-path \( P_i \) in \( H \) has length at most two for every \( i \in \{1, \ldots, k\} \); and

- for every \( i \in \{1, \ldots, k\} \), a set \( A_i \subseteq N(v_i) \) such that \( A_i \) is anticomplete to \( H \setminus \{v_i\} \) and \( A_j \) for \( j \neq i \), and such that \( |N(A_i)| \geq c|V(G)| \).

**Lemma 3.3.14.** Let \( t \in \mathbb{N} \) and \( C \geq 1 \). Then there exist \( d, \varepsilon > 0 \) with \( d \geq C\varepsilon \) such that for all \( c > 0 \) with \( c \leq 1/2 \), and for every \( \varepsilon \)-sparse graph \( G \) with no \( c \)-sparse \((\varepsilon cn, \varepsilon n)\)-pair, there is a \((t, d)\)-star cover in \( G \), where \( n = |V(G)| \).

**Proof.** By Lemma \ref{lemma:coherent_graph} for \( T = K_{1,t} \) and for \( v \) chosen as the center of \( T \), it follows that there exist \( c', \varepsilon' > 0 \) such that \( c' \geq 3C \cdot \varepsilon' \) and every \( \varepsilon' \)-coherent graph \( G' \) contains a \((T, v)\)-cover structure \( \{A_i\}_{i \in V(T)} \) with \( |A_v| \geq c'n \), where \( n = |V(G)| \).

Moreover, every \( A_t \) with \( t \in V(T) \setminus \{v\} \) is chosen according to Lemma \ref{lemma:coherent_graph} applied to a \((k', d')\)-box array with \( d' \geq c' \) and \( |N_G(A_t)| \geq c'n \).

We let \( \varepsilon' \) be as in Lemma \ref{lemma:coherent_graph} and let \( \varepsilon = \varepsilon'/4 \), and \( k = 1 \). Let \( G \) be an \( \varepsilon \)-sparse graph, \( c > 0 \), and suppose that \( G \) does not have a \( c \)-sparse \((\varepsilon cn, \varepsilon n)\)-pair. By Lemma \ref{lemma:large_neighbors} it follows that for every \( X \subseteq V(G) \) with \( |X| \geq 2\varepsilon'n \), \( X \) contains a vertex \( v \) with \( |N_G|X(v)| \geq c|X|/2 \), and since \( |X| \geq 2\varepsilon'n \), it follows that \( |N_G|X(v)| \geq \varepsilon cn \), and thus \( |V(G) \setminus N_G(N_G|X(v)|) \leq \varepsilon n \).

Therefore, when constructing \( \{A_t\}_{t \in V(T)} \) using Lemma \ref{lemma:coherent_graph} we may assume that for every \( t \in V(T) \setminus \{v\} \), \( A_t \) is chosen by applying Lemma \ref{lemma:coherent_graph} to a set \( B_i \) with \( |B_i| \geq c'n \geq 3\varepsilon n \). It follows that \( B_t \) contains a vertex \( w \) with \( |N_G|B_t(w)| \geq \varepsilon cn \).
Let $B'_i = N_{G[B_i]}[w]$ and let $\{a_1, \ldots, a_s\}$ be a labeling of $B'_i$ with $a_1 = w$. This implies that we may assume that every $A_t$ with $t \in V(T) \setminus \{v\}$ contains a vertex $v_t \in A_t$ such that $A_t = N_{G[A_t]}[v_t]$.

Now, by Lemma 3.2.14, we let $\{A_t\}_{t \in V(T)}$ be a $(T, v)$-cover structure such that

- For every $t \in V(T) \setminus \{v\}$, there exists a vertex $v_t \in A_t$ such that $A_t = N_{G[A_t]}[v_t]$, $|N_{G}(A_t)| \geq 3C\varepsilon n$ and $|A_t| \leq \varepsilon n$; and
- $|A_v| \geq 3C\varepsilon n$.

Let $V(T) = \{v, u_1, \ldots, u_t\}$. Let $w \in A_v$, and for $i \in \{1, \ldots, t\}$, let $w_i \in A_{u_i} \cap N(w)$; this exists since $A_{u_i}$ covers $A_v$. We construct a star $H$ with center $w$. For all $i \in \{1, \ldots, t\}$, we consider three cases:

- If $|N(A_{u_i} \cap N(w))| \geq C\varepsilon n$, we let $P_i = w$ and $A_i = A_{u_i} \cap N(w)$;
- If the first condition does not hold, and if $|N((A_{u_i} \cap N(w_i)) \setminus N(w))| \geq C\varepsilon n$, we let $P_i = w - w_i$ and $A_i = (A_{u_i} \cap N(w_i)) \setminus N(w)$;
- If the first and second conditions do not hold, and if $|N(A_{u_i} \setminus (N(w) \cup N(w_i)))| \geq C\varepsilon n$, we let $P_i = w - w_i - v_i$ and let $A_i = A_{u_i} \setminus (N(w) \cup N(w_i))$.

Since

$$A_{u_i} = (A_{u_i} \cap N(w)) \cup ((A_{u_i} \cap N(w_i)) \setminus N(w)) \cup (A_{u_i} \setminus (N(w) \cup N(w_i)))$$

and since $|N(A_{u_i})| \geq 3C\varepsilon n$, it follows that one of these cases occurs. We let $H = G|(V(P_1) \cup \ldots \cup V(P_t))$.

Since $A_{u_i}$ is anticomplete to $A_{u_j}$ for $i \neq j$, and since $V(P_i) \setminus \{w\} \subseteq A_{u_i}$, it follows that $H$, together with $A_1, \ldots, A_t$, is a $(t, d)$-star cover. This concludes the proof. □

**Lemma 3.3.15.** Let $\frac{1}{2} \geq \varepsilon > 0$, and let $G$ be an $\varepsilon$-sparse graph, $l \geq 0$, and let $X \subseteq V(G)$ with $|X| \geq (l + 4)\varepsilon n$. Let $y \in V(G) \setminus X$. Then one of the following statements holds:
(a) $G$ has an anticomplete $(\varepsilon n^\varepsilon, \varepsilon n)$-pair;

(b) $|N(y) \cap X| < \varepsilon n^\varepsilon$; or

(c) there is a path $P = v_0 - \ldots - v_l$ and $X' \subseteq X$ such that $v_0 = y$, $V(P) \setminus \{v_l\}$ is anticomplete to $X'$ and disjoint from $X'$, and $|N(v_l) \cap X'| \geq \varepsilon n^\varepsilon$;

and for all $c > 0$ with $c \leq 1/2$, one of the following statements holds:

(i) $G$ has a $c$-sparse $(\varepsilon cn, \varepsilon n)$-pair;

(ii) $|N(y) \cap X| < \varepsilon cn$; or

(iii) there is a path $P = v_0 - \ldots - v_l$ and $X' \subseteq X$ such that $v_0 = y$, $V(P) \setminus \{v_l\}$ is anticomplete to $X'$ and disjoint from $X'$, and $|N(v_l) \cap X'| \geq \varepsilon cn$.

Proof. We prove this by induction on $l$. If $l = 0$, then either $P = y$ and $X' = X$ satisfy (c), or $y$ satisfies (b); and either $P = y$ and $X' = X$ satisfy (iii), or $y$ satisfies (ii).

Now let $l > 0$. We first assume that $G$ has no anticomplete $(\varepsilon n^\varepsilon, \varepsilon n)$-pair; and that $|N(y) \cap X| \geq \varepsilon n^\varepsilon$. Let $A = N(y) \cap X$. It follows that $|X \setminus N[A]| \leq \varepsilon n$ since $|A| \leq \varepsilon n$ (because $G$ is $\varepsilon$-sparse) and $A, X \setminus N[A]$ is an anticomplete pair. Let $A' \subseteq A$ be minimal subject to $|N(A') \cap (X \setminus A)| \geq |X| - 3\varepsilon n$. By the minimality of $A'$, it follows that $|X \setminus (A \cup N(A'))| \geq |X \setminus A| - (|X| - 2\varepsilon n) \geq \varepsilon n$. Since $A', X \setminus (A \cup N(A'))$ is an anticomplete pair, it follows that $|A'| \leq \varepsilon n^\varepsilon$. It follows that there is a vertex $y' \in A'$ with

$$|N(y') \cap (X \setminus A)| \geq |N(A') \cap (X \setminus A)|/|A'| \geq (l + 1)\varepsilon n/\varepsilon n^\varepsilon = n^{1-\varepsilon} \geq \varepsilon n^\varepsilon.$$ 

Let $X' = X \setminus A$; it follows that $|X'| \geq (l + 3)\varepsilon n$. By induction applied to $y'$, $X'$ and $l - 1$, and since $|N(y') \cap X'| \geq \varepsilon n^\varepsilon$, it follows that either (a) holds, or there is a path $P' = v_0 - \ldots - v_{l-1}$ and $X'' \subseteq X'$ such that $v_0 = y'$, $V(P') \setminus \{v_{l-1}\}$ is anticomplete to $X''$. Therefore, by the induction hypothesis, either (a) holds, or there is a path $P'' = v_0 - \ldots - v_{l-1}$ and $X''' \subseteq X'$ such that $v_0 = y'$, $V(P'')$ is anticomplete to $X'''$. This completes the proof.
\(X''\) and disjoint from \(X''\), and \(|N(v_{l-1}) \cap X''| \geq \varepsilon n^\varepsilon\). We may assume that the latter is the case; now let \(P = y-y'P'v_{l-1}\). Since \(y\) is anticomplete to \(X \setminus A\) and disjoint from \(X \setminus A\), it follows that \(P\) and \(X''\) satisfy (c). This proves the first half of the statement.

Now let \(c > 0\), and suppose that \(G\) has no anticomplete \((\varepsilon cn, \varepsilon n)\)-pair; and that \(|N(y) \cap X| \geq \varepsilon cn\). Let \(A = N(y) \cap X\). Since \(G\) is \(\varepsilon\)-sparse, it follows that \(|X \setminus A| \geq (l + 3)\varepsilon n\). Since \(A, X \setminus A\) is not a \(c\)-sparse pair, it follows from Lemma 3.3.8 that there is a vertex \(y' \in A\) with \(|N(y') \cap (X \setminus A)| \geq \varepsilon cn\). Let \(X' = X \setminus A\); it follows that \(|X'| \geq (l+3)\varepsilon n\). By induction applied to \(y', X'\) and \(l-1\), and since \(|N(y') \cap X'| \geq \varepsilon cn\), it follows that either (i) holds, or there is a path \(P' = v_0 \ldots v_{l-1}\) and \(X'' \subseteq X'\) such that \(v_0 = y', V(P') \setminus \{v_{l-1}\}\) is anticomplete to \(X''\) and disjoint from \(X''\), and \(|N(v_{l-1}) \cap X''| \geq \varepsilon cn\). We may assume that the latter is the case; now let \(P = y-y'P'v_{l-1}\). Since \(y\) is anticomplete to \(X \setminus A\) and disjoint from \(X \setminus A\), it follows that \(P\) and \(X''\) satisfy (iii). This proves the second half of the statement. \(\square\)

Lemma 3.3.16. Let \(k, l \in \mathbb{N}\). Then there exists a \(C = C(k, l) \in \mathbb{N}\) such that the following holds. Let \(\varepsilon > 0\) with \(\varepsilon \leq 1/2C\), and let \(G\) be an \(\varepsilon\)-sparse graph. Let \(X_1, \ldots, X_k, Y_1, \ldots, Y_k \subseteq V(G)\) be pairwise disjoint and pairwise anticomplete and such that \(|N(X_i)|, |N(Y_i)| \geq C\varepsilon n\). Let \(l_1, \ldots, l_k \in \mathbb{N}\) with \(l_i \leq l\) for all \(i \in \{1, \ldots, k\}\). Then either there exist induced paths \(P_1, \ldots, P_k\) such that

- for all \(i \in \{1, \ldots, k\}\), \(P_i\) has length \(l_i\);
- for all \(i \in \{1, \ldots, k\}\), \(P_i\) has one end in \(X_i\) and one end in \(Y_i\), and the interior of \(P_i\) is disjoint from \(X_i\) and \(Y_i\); and
- for all \(i, j \in \{1, \ldots, k\}, i \neq j\), \(V(P_i)\) is anticomplete to \(X_j \cup V(P_j) \cup Y_j\),

or

- either there exists an \(i \in \{1, \ldots, k\}\) with \(l_i < 3\) or \(G\) has an anticomplete \((\varepsilon n^\varepsilon, \varepsilon n)\)-pair; and
either there exists an $i \in \{1, \ldots, k\}$ with $l_i < 4$ or for all $c > 0$ with $c \leq 1/2$, $G$

has a $c$-sparse $(\varepsilon cn, \varepsilon n)$-pair.

Proof. We prove this by induction on $k$. For $k = 0$, the result is trivially true with $C = 1$.

Now suppose that $k > 0$, and that the result holds for $k - 1$ and $l$ with constant $C(k - 1, l)$. We let $C = C(k, l) = 2k(C(k - 1, l) + l + 5)$. For $i \in \{1, \ldots, k - 1\}$, we let $X'_i \subseteq X_i$ be minimal with respect to inclusion subject to $|N(X'_i)| \geq (C(k - 1, l) + l + 4)\varepsilon n$, and we let $Y'_i \subseteq Y_i$ be minimal with respect to inclusion subject to $|N(Y'_i)| \geq (C(k - 1, l) + l + 4)\varepsilon n$. By the minimality of $X'_i$, it follows that $|N(X'_i)| \leq (C(k - 1, l) + l + 5)\varepsilon n$, and similarly, $|N(Y'_i)| \leq (C(k - 1, l) + l + 5)\varepsilon n$. We may assume that $X_k, Y_k$ are minimal with respect to inclusion subject to $|N(X_k)|, |N(Y_k)| \geq C(k, l)\varepsilon n \leq n/2$.

It follows that since $X_k, V(G) \setminus N[X_k]$ and $Y_k, V(G) \setminus N[Y_k]$ are anticomplete pairs, that we may assume that $|X_k|, |Y_k| \leq \varepsilon n$ (for otherwise the second condition in each of the last two bullets holds).

We now construct $P_k$. Let $Z = \bigcup_{i \in \{1, \ldots, k - 1\}} (N(Y'_i) \cup N(X'_i))$. It follows that $|Z| \leq 2(k - 1)(C(k - 1, l) + l + 5)\varepsilon n$. We let $A = N(X_k) \setminus Z$ and $B = N(Y_k) \setminus Z$. It follows that $|A|, |B| \geq 2(C(k - 1, l) + l + 4)\varepsilon n$.

Let $X'_k \subseteq X_k$ be minimal with respect to inclusion subject to $|N(X'_k) \cap (A \cup B)| \geq (l + 3)\varepsilon n$. By the minimality of $X'_k$, it follows that $|N(X'_k) \cap (A \cup B)| \leq (l + 4)\varepsilon n$. Let $C = B \setminus N(X'_k)$, and let $Y'_k = Y_k$. It follows that $|C| \geq (l + 4)\varepsilon n$. We will construct a path $P_k$ of length $l_k$ with ends in $X'_k, Y'_k$ and interior in $A \cup B$.

If $l_k \geq 3$ and $G$ has no anticomplete $(\varepsilon n^\varepsilon, \varepsilon n)$-pair, then there is a path $P_k$ with one end in $X'_k$, the other end in $Y'_k$, such that $P_k$ has length $l_k$ and the interior of $P_k$ is contained in $A \cup B$.

Suppose that $l_k \geq 3$ and $G$ has no anticomplete $(\varepsilon n^\varepsilon, \varepsilon n)$-pair. It follows that since $X_k, V(G) \setminus N[X_k]$ and $Y_k, V(G) \setminus N[Y_k]$ are anticomplete pairs, that $|X_k|, |Y_k| \leq \varepsilon n^\varepsilon$. 133
We will use Lemma 3.3.15 to construct a path $P$ of length $l_k - 3$ in $G$ such that $P$

starts at a vertex $y$ in $Y_k$, and $V(P) \setminus \{y\} \subseteq C$, and the last vertex of $P$ has at least

$\varepsilon n^\varepsilon$ neighbors in $C$. Since $Y_k$ covers $B$, and since $|Y_k| \leq \varepsilon n^\varepsilon$, it follows that there is a

vertex $y \in Y^k$ with $|N(y) \cap C| \geq (l + 4)n^{1-\varepsilon} \geq \varepsilon n^\varepsilon$. We now apply the first statement

of Lemma 3.3.15 to $y$, $C$, and $l_k - 3$. Since $G$ has no anticomplete $(\varepsilon n^\varepsilon, \varepsilon n)$-pair, and

since $|N(y) \cap C| \geq (l + 4)n^{1-\varepsilon} \geq \varepsilon n^\varepsilon$, it follows that outcome (c) of Lemma 3.3.15

occurs, and hence there is a path $P = v_0 - \ldots - v_{l_k-3}$ and $X' \subseteq C$ such that $v_0 = y$, $V(P) \setminus \{v_{l_k-3}\}$ is anticomplete to $X'$ and disjoint from $X'$, and $|N(v_{l_k-3}) \cap X'| \geq \varepsilon n^\varepsilon$.

Let $A' = A \setminus N[V(P)]$; it follows that $|A'| \geq 3\varepsilon n$. Since $N(v_{l_k-3}) \cap X', A'$ is not

c-sparse, it follows that there is a vertex $v_{l_k-2} \in N(v_{l_k-3}) \cap X'$ with a neighbor $v_{l_k-1} \in A'$. By the definition of $A'$, it follows that $v_{l_k-1}$ is anticomplete to $V(P)$. Let

$v_{l_k}$ be a neighbor of $v_{l_k-1}$ in $X'_k$. It follows that $v_{l_k}$ is anticomplete to $C$, and hence to

$v_0, \ldots, v_{l_k-2}$. We let $P_k = v_0 - v_1 - \ldots - v_{l_k-3} - v_{l_k-2} - v_{l_k-1} - v_{l_k}$. It follows that $P_k$ is a path

with ends in $X'_k, Y'_k$. Since the interior of $P_k$ is contained in $A \cup B$, it follows that $P_k$ is disjoint from $Z$ and hence $X'_j, Y'_j$ are anticomplete to $P_k$ for $j \in \{1, \ldots, k - 1\}$. This proves (3.9).

If $l_k \geq 4$ and there exists a $c > 0$ such that $G$ has no anticomplete

$(\varepsilon cn, \varepsilon n)$-pair, then there is a path $P_k$ with one end in $X'_k$, the other end

in $Y'_k$, such that $P_k$ has length $l_k$ and the interior of $P_k$ is contained in

$A \cup B$.

Suppose that $l_k \geq 4$ and $G$ has no anticomplete $(\varepsilon cn, \varepsilon n)$-pair. By Lemma 3.3.7

and since $|C| \geq (l + 4)\varepsilon n$, it follows that there is a vertex $y \in V(C)$ with $|N(y) \cap C| \geq

c|C|/2 \geq 2c\varepsilon n$. Let $y'$ be a neighbor of $y$ in $Y'_k$. If $|N(y') \cap C| \geq \varepsilon cn$, we let

$z = y'$; otherwise we let $z = y$. In the first case, we let $C' = C$; otherwise we let

$C' = C \setminus N(y')$. It follows that $|N(z) \cap C'| \geq \varepsilon cn$; in the first case by the choice of

134
This proves (3.10).

We apply the second statement of Lemma 3.3.15 to $z$, $C'$, and $l'$. Since $G$ has no anticomplete $(\varepsilon cn, \varepsilon n)$-pair, and since $|N(z) \cap C| \geq \varepsilon cn$, it follows that outcome (iii) occurs and there is a path $P = v_0 \ldots v_l'$ and $X' \subseteq C'$ such that $v_0 = z$, $V(P) \setminus \{v_l'\}$ is anticomplete to $X'$ and disjoint from $X'$, and $|N(v_l') \cap X'| \geq \varepsilon cn$.

Let $A' = A \setminus N[V(P) \cup \{y'\}]$; it follows that $|A'| \geq 3\varepsilon n$. Since $N(v_l') \cap X', A'$ is not $c$-sparse, it follows that there is a vertex $v_{l+1} \in N(v_{l'}) \cap X'$ with a neighbor $v_{l+2} \in A'$. By the definition of $A'$, it follows that $v_{l+2}$ is anticomplete to $V(P) \cup \{y'\}$. Let $v_{l+3}$ be a neighbor of $v_{l+2}$ in $X'_k$. It follows that $v_{l+3}$ is anticomplete to $C$, and hence to $v_0, \ldots, v_{l+1}$. We let $P_k = y'v_1\ldots v_{l'-1}v_{l+1}v_{l+2}v_{l+3}$ in the first case, and $P_k = y'y_{l'}v_{l'-1}\ldots v_{l'-2}v_{l+2}v_{l+3}$ in the second case. It follows that $P_k$ is a path of length $l_k$ with ends in $X'_k, Y'_k$. Since the interior of $P_k$ is contained in $A \cup B$, it follows that $P_k$ is disjoint from $Z$ and hence $X'_j, Y'_j$ are anticomplete to $P_k$ for $j \in \{1, \ldots, k-1\}$. This proves (3.10).

In (3.9) and (3.10), we have proved that either there is a path $P_k$ with one end in $X'_k$, the other end in $Y'_k$, such that $P_k$ has length $l_k$ and the interior of $P_k$ is contained in $A \cup B$, or

- either $l_k < 3$, or $G$ has an anticomplete $(\varepsilon n^c, \varepsilon n)$-pair; and
- either $l_k < 4$, or for all $c > 0$, $G$ has a $c$-sparse $(\varepsilon cn, \varepsilon n)$-pair.

Therefore, we may assume that we succeeded in constructing $P_k$. We let $G' = G \setminus (N[P_k])$. It follows that

$$|N(X'_i) \cap V(G')| \geq |N(X'_i)| - |V(P_k)| \varepsilon n$$

$$\geq (C(k - 1, l) + l + 4) \varepsilon n - (l + 1) \varepsilon n \geq C(k - 1, l) \varepsilon n.$$
It follows that $X'_1, \ldots, X'_k, Y'_1, \ldots, Y'_k \subseteq V(G')$ are pairwise disjoint and pairwise anticomplete and satisfy that $|N_{G'}(X'_i)|, |N_{G'}(Y'_i)| \geq C(k - 1, l)\varepsilon n$. By induction, it follows that either we can find $P_1, \ldots, P_{k-1}$ as desired, or

- either $l_k < 3$, or $G$ has an anticomplete $(\varepsilon n^c, \varepsilon n)$-pair; and
- either $l_k < 4$, or for all $c > 0$, $G$ has a $c$-sparse $(\varepsilon cn, \varepsilon n)$-pair.

Again, we may assume that we succeeded in constructing $P_1, \ldots, P_{k-1}$. By adding $P_k$, which is anticomplete to $V(G')$ and hence to $V(P_1), \ldots, V(P_{k-1})$, the result of the lemma follows.

The following result shows that unlike in Theorem 3.2.17, in this context every graph without the desired pair of sets contains every fixed subdivision of every graph as an induced subgraph, provided that the subdivision is sufficiently long:

**Theorem 3.3.17.**

- Let $H'$ be a graph, and let $H$ be a $(\geq 9)$-subdivision of $H'$. Then there exists $\varepsilon > 0$ such that if $G$ is $H$-free and $\varepsilon$-sparse, then $G$ contains an anticomplete $(\varepsilon n^c, \varepsilon n)$ pair, where $n = |V(G)|$.

- Let $H'$ be a graph, and let $H$ be a $(\geq 10)$-subdivision of $H'$. Then there exists $\varepsilon > 0$ such that for every $c > 0$ with $c \leq 1/2$, if $G$ is $H$-free and $\varepsilon$-sparse, then $G$ contains a $c$-sparse $(\varepsilon cn, \varepsilon n)$ pair, where $n = |V(G)|$.

**Proof.** We may assume that $H'$ is a complete graph, since for every $(\geq 9)$-subdivision $H^*$ of $H'$, there is a $(\geq 9)$-subdivision of $K_{|V(H')}|$ containing $H^*$ as an induced subgraph. We let $l$ be the maximum length of a path in $H$ corresponding to an edge of $H'$.

We let $t = |V(H')| - 1$, and $C = C(t(t+1)/2, l) + 3(t+1)/2$ with $C(t(t+1)/2, l)$ as in Lemma 3.3.16. By Lemma 3.3.14 there exist $d, \varepsilon' > 0$ with $d \geq C\varepsilon'$ such that for all $c > 0$, and for every $\varepsilon'$-sparse graph $G$ with no $c$-sparse $(\varepsilon' cn, \varepsilon'n)$-pair, there is a $(t, d)$-star cover in $G$, where $n = |V(G)|$. 

136
Now let $\varepsilon = \varepsilon'/(2t^2(C + 4))$ let $G$ be an $\varepsilon$-sparse graph, and suppose that there exists $c > 0$ such that $G$ contains no $c$-sparse $(\varepsilon cn, \varepsilon n)$ pair, where $n = |V(G)|$.

We let $V_0 = V(G)$ and repeat the following for $i \in \{1, \ldots, t + 1\}$: Since $|V_{i-1}| \geq n - t(i-1)(C+4)\varepsilon n \geq n/2$ and $\varepsilon \leq \varepsilon'/2$, it follows that by Lemma 3.3.14 $G|_{V_{i-1}}$ contains a $(t, C\varepsilon')$-star cover $J^i, B_1^i, \ldots, B_t^i$. Since $\varepsilon' \geq 2\varepsilon$, it follows that $J^i, B_1^i, \ldots, B_t^i$ is a $(t, C\varepsilon)$-star cover in $G$. For $j \in \{1, \ldots, t\}$, we let $A_j^i \subseteq B_j^i$ be minimal subject to $|N_G(A_j^i)| \geq C\varepsilon n$. By the minimality of $A_j^i$, it follows that $|N(A_j^i)| \leq (C + 1)\varepsilon n$ for all $j \in \{1, \ldots, t\}$. Therefore, it follows that $|N[V(J^i) \cup A_1^i \cup \cdots \cup A_t^i]| \leq t(C + 4)\varepsilon n$. We let $V_i = V_{i-1} \setminus N[V(J^i) \cup A_1^i \cup \cdots \cup A_t^i]$. It follows that $|V_i| \geq n - ti(C + 4)\varepsilon n$.

We let $C^i = J^i, A_1^i, \ldots, A_t^i$ denote the $(t, C\varepsilon)$-star cover we constructed.

This defines a sequence of $(t, d)$-star covers $J^1, A_1^1, \ldots, J^{t+1}, A_1^{t+1}, \ldots, A_t^{t+1}$ with the additional property that for all $i, j \in \{1, \ldots, t + 1\}$ with $i \neq j$, $V(J^i) \cup A_1^i \cdots \cup A_t^i$ is anticomplete to $V(J^j) \cup A_1^j \cdots \cup A_t^j$.

Now let $e_1, \ldots, e_s$ be the edges of $H' = K_{|V(H')|}$, let $V(H') = \{v_1, \ldots, v_{t+1}\}$ and let $f$ be a function that maps a pair $(i, j)$ with $i \in \{1, \ldots, t+1\}$ and $j \in \{1, \ldots, t\}$ to an edge $e = uv$ with $v_i \in \{u, v\}$, and such that for all $e \in E(H')$, $f^{-1}(e) = \{(i_1, j_1), (i_2, j_2)\}$ and $e = v_{i_1}v_{i_2}$. For $p \in \{1, \ldots, s\}$, we let $f^{-1}(e_p) = \{(i_1, j_1), (i_2, j_2)\}$, and let $X_p = A_{j_1}^{i_1}$ and $Y_p = A_{j_2}^{i_2}$. We let

$$G' = G \setminus \left( N[V(J^1 \cup \cdots \cup V(J^{t+1})]) \setminus (X_1 \cup \cdots \cup X_s \cup Y_1 \cup \cdots \cup Y_s) \right).$$

It follows that $|N_{G'}(X_i)| \geq |N_G(X_i)| - |V(J^1 \cup \cdots \cup V(J^{t+1}))-\varepsilon n \geq (C(t+1)/2, l) + 3t(t+1)/2 - 3t(t+1)/2\varepsilon n = C(t+1)/2, l)\varepsilon n$, and similarly, $|N_{G'}(Y_i)| \geq C(t+1)/2, l)\varepsilon n$. Moreover, the sets $X_1, \ldots, X_s, Y_1, \ldots, Y_s$ are pairwise disjoint and pairwise anticomplete by construction, and $s = |E(H')| = t(t+1)/2$. For $i \in \{1, \ldots, s\}$, we let $L_i$ denote the length of the path in $H$ corresponding to the edge $e_i \in E(H')$. Let $f^{-1}(e_i) = \{(i_1, j_1), (i_2, j_2)\}$; and let $x^i_1$ denote the length of the path from the center of
to the vertex \( x \) of \( J^i \) with \( A^i_{j_1} \subseteq N(x) \); let \( x_2^i \) denote the length of the path from the center of \( J^{i_2} \) to the vertex \( x' \) of \( J^{i_2} \) with \( A^{i_2}_{j_2} \subseteq N(x) \). We let \( l_i = L_i - x^i_1 - x^i_2 - 2 \).

Since \( x_1 \leq 2, x_2 \leq 2 \) by the definition of a \((t,d)\)-star cover, it follows that \( l_i \geq L_i - 6 \).

Now we apply Lemma 3.3.10 to \( X_1, \ldots, X_s, Y_1, \ldots, Y_s \) and \( l_1, \ldots, l_s \). Since \( H \) is a \((\geq 9)\)-subdivision of \( H' \) for the first bullet, and a \((\geq 10)\)-subdivision of \( H' \) for the second bullet, it follows that \( l_i \geq 3 \) for all \( i \in \{1, \ldots, s\} \) for the first bullet, and \( l_i \geq 4 \) for all \( i \in \{1, \ldots, s\} \) for the second bullet; therefore we obtain paths \( P_1, \ldots, P_s \) such that

- for all \( i \in \{1, \ldots, s\} \), \( P_i \) has length \( l_i \);
- for all \( i \in \{1, \ldots, s\} \), \( P_i \) has one end in \( X_i \) and one end in \( Y_i \), and the interior of \( P_i \) is disjoint from \( X_i \) and \( Y_i \); and
- for all \( i, j \in \{1, \ldots, s\}, i \neq j \), \( V(P_i) \) is anticomplete to \( X_j \cup V(P_j) \cup Y_j \).

Finally, we are ready to exhibit a copy of \( H \) in \( G \). For \( i \in \{1, \ldots, s\} \), the vertex \( v_i \in V(H') \) corresponds to the center of \( J^i \). For every edge \( e_s \in E(H') \) with \( f^{-1}(e_i) = \{(i_1, j_1), (i_2, j_2)\} \), we let \( Q_s \) be the path obtained by concatenating the path from the center of \( J^{i_1} \) to the vertex \( x \) of \( J^{i_1} \) with \( X_s = A^{i_1}_{j_1} \subseteq N(x) \) with \( P_s \) and with the path from the vertex \( x' \) of \( J^{i_2} \) with \( Y_s = A^{i_2}_{j_2} \subseteq N(x) \) to the center of \( J^{i_2} \). It follows that \( Q_s \) has length \( l_s + 2 + x^i_1 + x^i_2 = L_i \). Therefore \( G|\{V(J^1) \cup \cdots \cup V(J^{t+1}) \cup V(P_1) \cup \cdots \cup V(P_s)\} \) is isomorphic to \( H \), as claimed.

\[ \square \]

3.4 Polynomial complete or anticomplete pairs

In this section, we adapt the proof strategy used to prove the following two results, the latter of which was used to prove Theorem 3.1.6.
Theorem 3.4.1 (Erdős, Hajnal, Pach [23]). For every $k$-vertex graph $H$ and for every $H$-free graph $G$, $G$ contains a complete or anticomplete $((n/k)^{1/(k-1)}, (n/k)^{1/(k-1)})$-pair, where $n = |V(G)|$.

Theorem 3.4.2 (Erdős, Hajnal [22]). For every graph $H$ there is a constant $k$ such that for every $c$ with $1/2 \geq c > 0$ and for every $H$-free graph $G$, $G$ contains either a $c$-sparse or a $(1-c)$-dense $(cn^{k+1}c, cn^{1-kc})$-pair, where $n = |V(G)|$.

The following result is a variant of Theorem 3.4.1:

Theorem 3.4.3. Let $H$ be a graph. Then there exist constants $c, \varepsilon > 0$ such that every $H$-free graph $G$ with $|V(G)| = n \geq |V(H)|$ contains a complete or anticomplete $(cn^{k+1}\varepsilon, cn^{1-kc})$-pair for some $k \in \{2j - 1 : j \in \{0, \ldots, |V(H)| - 2\}\}$.

Proof. We may assume that $|V(H)| \geq 2$; we let $V(H) = \{v_1, \ldots, v_t\}$, $d = 1/(2t)$, $c = d^2$, and $\varepsilon = 2^{-t}$.

Let $G$ be a graph, $|V(G)| = n$. We may assume that for $k = 2^{t-2} - 1$, we have $cn^{k+1}\varepsilon > 1$, for otherwise, for every two distinct vertices $u \in V(G)$, either $\{u\}, N(u)$ or $\{u\}, V(G) \setminus N[u]$ is a complete or anticomplete $(cn^{k+1}\varepsilon, cn)$-pair. We partition $V(G)$ into $t$ sets $A_1, \ldots, A_t$, each of size at least $|V(G)|/2t \geq dn$. We will prove that either $G$ contains a copy of $H$ such that the vertex corresponding to $v_i$ is contained in $A_i$ for all $i$, or $G$ contains a complete or anticomplete $(cn^{k+1}\varepsilon, cn^{1-kc})$-pair for some $k \in \{2j - 1 : j \in \{0, \ldots, t\}\}$.

We call $x_1, \ldots, x_s, A_{s+1}, \ldots, A_t$ an $s$-partial embedding of $H$ if

- for all $i \in \{s + 1, \ldots, t\}$, $|A_i| \geq dn^{1-((2^{s-1}-1)c)}$;
- for all $i, j \in \{1, \ldots, s\}$, $x_ix_j \in E(G)$ if and only if $v_iv_j \in E(H)$; and
- for all $i \in \{1, \ldots, s\}, j \in \{s + 1, \ldots, t\}$, if $v_iv_j \in E(H)$, then $x_i$ is complete to $A_j$; otherwise $x_i$ is anticomplete to $A_j$. 

139
Clearly, if $G$ has a $t$-partial embedding of $H$, then $G$ contains $H$ as an induced subgraph. Moreover, $A_1, \ldots, A_t$ is a 1-partial embedding of $H$.

If $G$ has an $s$-partial embedding of $H$ with $s < t$, then either $G$ has an $(s + 1)$-partial embedding of $H$, or $G$ has a complete or anticomplete $(cn^{2s-1} \varepsilon, cn^{1-(2s-1)\varepsilon})$-pair.

Let $x_1, \ldots, x_s, A_{s+1}, \ldots, A_t$ be an $s$-partial embedding of $H$. For $j \in \{s + 2, \ldots, t\}$, we let $A_j$ be defined as follows:

- if $v_{s+1}v_j \in E(H)$, we let $A_j = \{a \in A_{s+1} : |N(a) \cap A_j| \leq dn^{1-(2s-1)\varepsilon}\}$; and
- if $v_{s+1}v_j \not\in E(H)$, we let $A_j = \{a \in A_{s+1} : |A_j \setminus N(a)| \leq dn^{1-(2s-1)\varepsilon}\}$.

Suppose that there is a $j \in \{s + 2, \ldots, t\}$ with $|A_j| \geq cn^{2s-1} \varepsilon$. By symmetry, we may assume that $v_{s+1}v_j \in E(H)$. Let $B \subseteq A_j$ with $|B| = \lceil cn^{2s-1} \varepsilon \rceil$. Since $cn^{2s-1} \varepsilon > 1$ by our assumption, it follows that $|B| \leq 2cn^{2s-1} \varepsilon$. Let $C = N(B) \cap A_j$. It follows that

$$|C| \leq 2cn^{2s-1} \varepsilon \cdot dn^{1-(2s-1)\varepsilon} \geq 2cdn^{1-(2s-1)\varepsilon} \leq dn^{1-(2s-1)\varepsilon}/2 \leq |A_j|/2,$$

and therefore $B, A_j \setminus C$ is an anticomplete $(cn^{2s-1} \varepsilon, cn^{1-(2s-1)\varepsilon})$-pair.

It follows that we may assume that $|A_j| \leq cn^{2s-1} \varepsilon$, and thus $|A_{s+1} \cup \ldots \cup A'_t| \leq tcn^{2s-1} \varepsilon \leq tcn^{1-(2s-1)\varepsilon} \leq (d/2)n^{1-(2s-1)\varepsilon} \leq |A_j|/2$. It follows that there exists a vertex $x_{s+1} \in A_j \setminus (A_{s+1} \cup \ldots \cup A'_t)$. For all $j \in \{s + 2, \ldots, t\}$, we let $A'_j = A_j \cap N(x_{s+1})$ if $v_{s+1}v_j \in E(H)$, and $A'_j = A_j \setminus N(x_{s+1})$ otherwise. It follows that $|A'_j| \geq dn^{1-(2s-1)\varepsilon}$, and so $x_1, \ldots, x_{s+1}, A'_{s+2}, \ldots, A'_t$ is an $(s + 1)$-partial embedding of $H$. This proves \((3.11)\).

By repeated application of \((3.11)\), it follows that either $G$ contains $H$ as an induced subgraph, or $G$ contains a complete or anticomplete $(cn^{(k+1)\varepsilon}, cn^{1-k\varepsilon})$-pair for some $k \in \{2^s - 1 : s \in \{0, \ldots, |V(H)| - 2\}\}$. This concludes the proof. \qed
The following is a simple consequence of Theorem 3.4.3, and it proves that a weakening of the property in Question 3.3.1 holds for every graph $H$.

**Corollary 3.4.4.** Let $H$ be a graph. Then there exist constants $c, \varepsilon > 0$ such that every $H$-free graph $G$ with $|V(G)| = n \geq |V(H)|$ contains a complete or anticomplete pair $A, B$ with $|A||B| \geq cn^{1+\varepsilon}$.

**Theorem 3.4.5.** Let $H$ be a graph. Then there exist constants $c, \varepsilon > 0$ such that for every $\delta > 0$ with $\delta \leq 1/2$, and for every $H$-free graph $G$ with $|V(G)| = n \geq |V(H)|$, the graph $G$ contains a $\delta^{k+1}$-sparse or $(1 - \delta^{k+1})$-dense $(c\delta^k n, c\delta^k n)$-pair for some $k \in \{2^j - 1 : j \in \{0, \ldots, |V(H)|\}\}$.

**Proof.** The proof is similar to the proof of Theorem 3.4.3. We may assume that $|V(H)| \geq 2$; we let $V(H) = \{v_1, \ldots, v_t\}$, $d = 1/(2t)$, $c = d^2$, and $\varepsilon = 2^{-t}$.

Let $G$ be a graph, $|V(G)| = n$, and let $\delta > 0$. We may assume that for $k = 2^{t-2} - 1$, we have that $c\delta^k n > 1$, for otherwise every two distinct vertices $u, v \in V(G)$ form a complete or anticomplete $(c\delta^k n, c\delta^k n)$-pair. We partition $V(G)$ into $t$ sets $A_1, \ldots, A_t$, each of size at least $|V(G)|/(2t) \geq dn$. We will prove that either $G$ contains a copy of $H$ such that the vertex corresponding to $v_i$ is contained in $A_i$ for all $i$, or $G$ contains a complete or anticomplete $(c\delta^k n, c\delta^k n)$-pair for some $k \in \{2^j - 1 : j \in \{0, \ldots, t\}\}$.

We call $x_1, \ldots, x_s, A_{s+1}, \ldots, A_t$ an $s$-partial embedding of $H$ if

- for all $i \in \{s+1, \ldots, t\}$, $|A_i| \geq d\delta^{2s-1-1}n$;
- for all $i, j \in \{1, \ldots, s\}$, if $x_i x_j \in E(G)$ if and only if $v_i v_j \in E(H)$; and
- for all $i \in \{1, \ldots, s\}$, $j \in \{s+1, \ldots, t\}$, if $v_i v_j \in E(H)$, then $x_i$ is complete to $A_j$; otherwise $x_i$ is anticomplete to $A_j$.
Clearly, if $G$ has a $t$-partial embedding of $H$, then $G$ contains $H$ as an induced subgraph. Moreover, $A_1, \ldots, A_t$ is a 1-partial embedding of $H$.

If $G$ has an $s$-partial embedding of $H$ with $s < t$, then either $G$ has an $(s + 1)$-partial embedding of $H$, or $G$ has a complete or anticomplete $(c\delta^{2s-1}n, c\delta^{2s-1}n)$-pair.

Let $x_1, \ldots, x_s, A_{s+1}, \ldots, A_t$ be an $s$-partial embedding of $H$. For $j \in \{s + 2, \ldots, t\}$, we let $A^j$ be defined as follows:

- if $v_{s+1}v_j \in E(H)$, we let $A^j = \{a \in A_{s+1} : |N(a) \cap A_s| \leq \delta^{2s-1}|A_j|\}$; and
- if $v_{s+1}v_j \not\in E(H)$, we let $A^j = \{a \in A_{s+1} : |A_j \setminus N(a)| \leq \delta^{2s-1}|A_j|\}$.

Suppose that there is a $j \in \{s + 2, \ldots, t\}$ with $|A^j| \geq c\delta^{2s-1}n$. By symmetry, we may assume that $v_{s+1}v_j \in E(H)$. Let $B \subseteq A^j$ with $|B| = \left\lceil c\delta^{2s-1}n \right\rceil$. Since $c\delta^{2s-1}n > 1$, it follows that $|B| \leq 2c\delta^{2s-1}n$. Let $C = A_j$. It follows that $|E(B, C)| \leq \delta^{2s-1}|B||C|$, and therefore $B, C$ is a $\delta^{2s-1}$-sparse $(c\delta^{2s-1}n, c\delta^{2s-1}n)$-pair.

It follows that we may assume that $|A^j| \leq c\delta^{2s-1}n$, and thus $|A_{s+1} \cup \cdots \cup A^j| \leq tc\delta^{2s-1}n \leq (d/2)\delta^{2s-1}n \leq |A_j|/2$. It follows that there exists a vertex $x_{s+1} \in A_j \setminus (A_{s+1} \cup \cdots \cup A^j)$. For all $j \in \{s + 2, \ldots, t\}$, we let $A'_j = A_j \cap N(x_{s+1})$ if $v_{s+1}v_j \in E(H)$, and $A'_j = A_j \setminus N(x_{s+1})$ otherwise. It follows that $|A'_j| \geq \delta^{2s-1}|A_j| \geq d\delta^{2s-1}n$, and so $x_1, \ldots, x_{s+1}, A'_{s+2}, \ldots, A'_t$ is an $(s + 1)$-partial embedding of $H$. This proves (3.11).

By repeated application of (3.11), it follows that either $G$ contains $H$ as an induced subgraph, or $G$ contains a complete or anticomplete $(c\delta^k n, c\delta^k n)$-pair for some $k \in \{2s - 1 : s \in \{0, \ldots, |V(H)| - 2\}\}$. This concludes the proof.
3.5 Sparse sets in sparse graphs

In this section, we extend results in Sections 3.3 and 3.4 by strengthening them in one or more of the following ways:

- Instead of excluding a graph $H$ as an induced subgraph of $G$, we bound the number of induced subgraphs of $G$ isomorphic to $H$.

- Instead of allowing both sparse and dense (complete and anticomplete) pairs as outcomes, if $G$ is sparse, we only allow sparse (anticomplete) pairs.

- Instead of two sets with a given edge density, we obtain $k$ sets with a given edge density.

In Section 3.5.1, we introduce a “$k$-tuple game”, which serves as our main tool for finding $k$ sets with a given edge density instead of two sets.

For a graph $H$ and a graph $G$, a copy of $H$ in $G$ is a pair $(T, \varphi)$ consisting of a subset $T \subseteq V(G)$ together with an isomorphism $\varphi$ between $H$ and $G|T$. In particular, $G$ contains $H$ as an induced subgraph if and only if there is a copy of $H$ in $G$. A graph $G$ is $(\alpha, H)$-saturated if $G$ contains $\alpha|V(G)||V(H)|$ distinct copies of $H$.

We use the following strengthening of Corollary 3.1.10 to reduce this problem to the sparse case:

**Theorem 3.5.1** (Fox and Sudakov [27]). Let $H$ be a graph, and let $\varepsilon > 0$. Then there exist $\alpha, \delta > 0$ such that for every graph $G$, if $G$ is not $(\alpha, H)$-saturated, then $G$ contains an induced subgraph $J$ with $|V(J)| \geq \delta|V(G)|$ and such that either $|E(J)| \leq \varepsilon|V(J)|(|V(J)| - 1)/2$ or $|E(J^c)| \leq \varepsilon|V(J)|(|V(J)| - 1)/2$.

The following corollary has the the same proof as for Corollary 3.1.10 (using Theorem 3.5.1 instead of Theorem 3.1.9).

**Corollary 3.5.2.** For every graph $H$ and every $\varepsilon > 0$ there exist $\alpha, \delta' > 0$ such that for every graph $G$, if $G$ is not $(\alpha, H)$-saturated, then $G$ either satisfies $|V(G)| \leq 1/\delta'$.
or contains an induced subgraph $J$ with $|V(J)| \geq \delta'|V(G)|$ and such that either $J$ or $J^c$ is $\varepsilon$-sparse.

As in Section 3.1.1, this allows us to consider the sparse case separately.

### 3.5.1 A game on a graph

For $k \geq 2$ and $m \geq k$, the $k$-tuple game on $m$ vertices is the following game between two players, $A$ and $B$. We let $G_0$ be a graph with $m$ vertices and no edges. During the $i$th round (starting with $i = 1$), the following moves happen in this order.

- $G_{i-1}$ is revealed to player $A$;
- player $A$ may decide to stop the game and output $G_{i-1}$;
- player $A$ chooses a $k$-vertex subset $T \subseteq V(G_0)$ such that $T$ is a stable set in $G_{i-1}$ (if there is no such set then player $A$ has to stop the game and output $G_{i-1}$);
- player $B$ chooses a graph $G_i$, as follows: $V(G_i) = V(G_{i-1})$, and $G_i$ arises from $G_{i-1}$ by adding edges with both ends in $T$ such that at least one edge was added.

The graph $G_i$ is then revealed to player $A$, and a new round commences. Since an edge is added in every round, this game terminates after a finite number of rounds.

For a graph $H$, we say that $H$ is $(m, k)$-forcible if there is a strategy for player $A$ to play the $k$-tuple game on $m$ vertices and always output a graph $G$ that contains $H$ as an induced subgraph. We say that such a strategy forces $H$.

The main result of this section is that for every $H$ and $k$, there exists an $m$ such that $H$ is $(m, k)$-forcible. We begin by proving the base cases in the next lemma.

**Lemma 3.5.3.** Let $H$ be a graph. If $|V(H)| = l$, then $H$ is $(l, 2)$-forcible. If furthermore $|E(H)| = 0$, then $H$ is $(l, k)$-forcible for all $k \geq 2$. 

144
Proof. The second statement holds since $G_0$ is isomorphic to $H$ if $H$ has no edges. For the first statement, player $A$ picks a bijection $f$ between $V(H)$ and $V(G_0)$, and player $A$ ensures that in every round $i$, $G_{i-1}$ is isomorphic to a (not necessarily induced) subgraph of $H$. Therefore either $H$ is isomorphic to $G_{i-1}$ or there is an edge $uv \in E(H)$ such that $f(u)f(v) \notin E(G_{i-1})$. In the former case, player $A$ stops the game; in the latter case, player $A$ picks the set $\{f(u), f(v)\}$. This forces player $B$ to add the edge $f(u)f(v)$. After $|E(H)|$ rounds, $G_i$ has $|E(H)|$ edges and hence is isomorphic to $H$. \qed

Lemma 3.5.4. Let $H_1, H_2$ be graphs, $k \geq 2$ and $m_1, m_2 \in \mathbb{N}$ such that for all $i \in \{1, 2\}$, $H_i$ is $(m_i, k)$-forcible. Then the disjoint union of $H_1$ and $H_2$ is $(m_1 + m_2, k)$-forcible.

Proof. To prove this, we let $G_0$ have $m_1 + m_2$ vertices partitioned into $V_1, V_2$ with $|V_i| = m_i$ for all $i \in \{1, 2\}$. Player $A$ first plays according to the the $k$-tuple game for $H_1$ on $G_0|V_1$. Since $H_1$ is $(m_1, k)$-forcible, this game stops after a finite number $s_1$ of rounds and $G_{s_1-1}|V_1$ contains $H_1$ as an induced subgraph. Since every $k$-tuple picked by player $A$ is contained in $V_1$, it follows that $V_2$ is stable and anticomplete to $V_1$ in $G_{s_1-1}$. Instead of stopping in round $s_1$, player $A$ now plays according to the $k$-tuple game for $H_2$ on $G_{s_1-1}|V_2$. After a finite number $s_2$ of rounds, $G_{s_1+s_2-1}|V_2$ contains $H_2$ as an induced subgraph, and $V_1$ remains anticomplete to $V_2$. But this implies that $G_{s_1+s_2-1}$ contains the disjoint union of $H_1$ and $H_2$ as an induced subgraph. \qed

Lemma 3.5.5. Let $H$ be a graph and $k \geq 2$. Then there exists an $m \geq k$ such that $H$ is $(m, k)$-forcible.

Proof. We prove this by induction on $k + |E(H)|$. The statement holds in the base cases, when either $k = 2$ or $|E(H)| = 0$, by Lemma 3.5.3. Now let $k > 2$ and $|E(H)| > 0$; let $e = uv \in E(H)$ and suppose that $H \setminus \{e\}$ is $(m_1, k)$-forcible and that $H$ is $(m_2, k-1)$-forcible. Let $m = m_1m_2$. We claim that $H$ is $(m, k)$-forcible.
A \((u, s)\)-star of a graph \(H'\) with \(u \in V(H)\) is a graph \(J\) with \(s(|V(H')| - 1) + 1\) vertices partitioned into sets \(V_1, \ldots, V_s, \{w\}\) such that

- \(|V_i| = |V(H')| - 1\) for all \(i \in \{1, \ldots, s\}\);
- \(V_i\) is anticomplete to \(V_j\) for all \(i, j \in \{1, \ldots, s\}\) with \(i \neq j\); and
- \(J|(V_i \cup \{w\})\) is isomorphic to \(H'\) and \(w\) corresponds to \(u\) for this isomorphism, for all \(i \in \{1, \ldots, s\}\).

The vertex \(w\) is called the center of \(H'\).

\[(3.13)\] For every \(s \geq 1\), a \((u, s)\)-star of \(H \setminus \{e\}\) is \((m^s_1, k)\)-forcible.

We prove this by induction on \(s\). For \(s = 1\), this follows since \(H \setminus \{e\}\) is \((m_1, k)\)-forcible and isomorphic to a \((u, 1)\)-star. Now let \(s > 1\). By the inductive hypothesis, a \((u, s - 1)\)-star is \((m^{s-1}_1, k)\)-forcible, and so by Lemma \[3.5.4\] the disjoint union of \(m_1\) graphs \(H_1, \ldots, H_{m_1}\), each isomorphic to a \((u, s - 1)\)-star, is \((m^s_1, k)\)-forcible. For \(i \in \{1, \ldots, m_1\}\), let \(u_i\) denote the center of \(H_i\). We let \(U = \{u_i : i \in \{1, \ldots, m_1\}\}\). It follows that \(|U| = m_1\) and \(U\) is stable. By applying the strategy for \(H \setminus \{e\}\) to \(U\), it follows that there is a strategy for player \(A\) that produces in some round \(p\) a graph \(G_p\) containing an induced subgraph that consists of \(m_1\) disjoint \((u, s - 1)\)-stars, a set \(U\) as defined above, \(H \setminus \{e\}\) as an induced subgraph with vertex set \(V_s\) contained in \(U\), and such that all edges between distinct copies of \((u, s - 1)\)-stars have both ends in \(U\). Let \(i \in \{1, \ldots, m_1\}\) such that \(u_i \in V_s\) corresponds to \(u\) in the isomorphism between \(G_p|V_s\) and the graph \(H \setminus \{e\}\). It follows that \(V_s \setminus \{u_i\}\) is anticomplete to \(V(H_i) \setminus \{u_i\}\), and therefore \(G_p|(V_s \cup V(H_i))\) is a \((u, s)\)-star. This implies \((3.13)\).

By \((3.13)\), it follows that a \((u, m_2)\)-star is \((m^{m_2}_1, k)\)-forcible. This implies that in the \(k\)-tuple game on \(m\) vertices, player \(A\) can guarantee that in some round \(s\), \(G_s\)
contains a \((u, m_2)\)-star \(H'\) with vertex set \(V_1 \cup \cdots \cup V_{m_2} \cup \{w\}\) as an induced subgraph. Let \(w\) be the center of \(H'\), and for \(i \in \{1, \ldots, m_2\}\), let \(v_i\) be the vertex corresponding to \(v\) in the isomorphism between \(G_s|(V_i \cup \{w\})\) and \(H \setminus \{e\}\). Let \(V = \{v_1, \ldots, v_{m_2}\}\). By the definition of a \((u, m_2)\)-star, it follows that \(V\) is a stable set in \(G_s\), and \(|V| = m_2\).

Now player \(A\) uses the strategy that forces \(H\) in the \((k - 1)\)-tuple game on \(m_2\) vertices by starting with \(G_s|V\); except in every round, player \(A\) picks \(T \cup \{w\}\) instead of the set \(T \subseteq V\) of \((k - 1)\) vertices that the strategy for the \((k - 1)\)-tuple game produces. If in round \(s' > s\), player \(B\) adds an edge incident with \(w\), say \(wv_i\), then \(G_{s'}|(V_i \cup \{w\})\) is isomorphic to \(H\), and the result follows. Therefore, we may assume that in every round \(s' > s\), player \(A\) picks a set \(T \cup \{w\}\), and player \(B\) adds at least one edge with both endpoints in \(T\). It follows that since \(H\) is \((m_2, k - 1)\)-forcible, for some \(s' > s\), \(G_{s'}|V\) contains \(H\) as an induced subgraph. This concludes the proof. \(\Box\)

Lemma 3.5.5 implies that player \(A\) can always win the \(k\)-game, given a sufficiently large starting graph. This motivates the following questions.

**Question 3.5.6.** Can player \(A\) always win the \(k\)-tuple game on a sufficiently large starting graph if in round \(i\), \(G_{i-1}\) is not revealed to player \(A\)?

This is equivalent to requiring player \(A\) to pick the vertex set for every round in advance, without depending on the choices made by player \(B\). There are a few variants; a win for player \(A\) could be defined as producing a given graph \(H\) as an induced subgraph at some point during the game, or at termination; also, if player \(A\) chooses a set that is not stable, this could either mean that player \(A\) loses, or that player \(B\) skips this set (adds no edges).
3.5.2 Many pairwise sparse pairs

In this section we prove that for every graph $H$ and $k \in \mathbb{N}$, there exists an $\varepsilon > 0$ such that every $\varepsilon$-sparse graph contains either $H$ or a sequence $A_1, \ldots, A_k$ such that $|A_i| \geq \varepsilon n^k$ and $A_i$ is anticomplete to $A_j$ for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$.

Lemma 3.5.7. Let $G$ be a graph, and let $H$ be a graph, and $\alpha > 0$ such that $G$ is $(\alpha, H)$-saturated. Let $H'$ be an induced subgraph of $H$. Then $G$ is $(\alpha, H')$-saturated.

Proof. Let $n = |V(G)|$. Let $\mathcal{T}$ be the set of copies of $H'$ in $G$. For every copy $(T, \varphi)$ of $H$ in $G$, $(\varphi(V(H')), \varphi|_{V(H')})$ is a copy of $H'$ in $G$; we say that $(T, \varphi)$ corresponds to $(\varphi(V(H')), \varphi|_{V(H')})$. For every $(T', \varphi') \in \mathcal{T}$, there are at most $n^{|V(H)| - |V(H')|}$ copies of $H$ that correspond to $(T', \varphi')$ (since there are at most $n^{|V(H)| - |V(H')|}$ ways to extend $\varphi'$ from a function from $V(H')$ to $V(G)$, to a function from $V(H)$ to $V(G)$). It follows that $|\mathcal{T}| \geq n^{|V(H')| - |V(H')|} \cdot n^{|V(H)|} = \alpha n^{|V(H')|}$, and so $G$ is $(\alpha, H')$-saturated.

Let $H$ be a graph, and let $S \subseteq V(H)$ be a stable set. Then a graph $H'$ is an $S$-successor of $H$ if

- $V(H) = V(H')$;
- $|E(H)| < |E(H')|$ and $E(H) \subseteq E(H')$; and
- for every $e \in E(H') \setminus E(H)$, we have $e = uv$ with $u, v \in S$.

In the $k$-tuple game, if player $A$ selects $S \subseteq H$ with $|S| = k$ and $S$ stable, then a graph $H'$ is a valid response by player $B$ if and only if $H'$ is an $S$-successor of $H$.

We begin with the following observation.

Lemma 3.5.8. Let $H$ be a graph, and let $S \subseteq V(H)$ be a stable set. Then there are at most $2^{|S|^2}$ distinct $S$-successors of $H$.

Proof. Every $S$-successor $H'$ is completely determined by $H'|S$, and there are at most $2^{|S|^2}$ distinct graphs on $|S|$ vertices.
Let $G$ be a graph, and let $A_1, \ldots, A_k \subseteq V(G)$ be pairwise disjoint such that

- for all $i \in \{1, \ldots, k\}$, $|A_i| \geq \alpha n$; and
- for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $A_i, A_j$ is a $c$-sparse pair.

Then we call $A_1, \ldots, A_k$ a $c$-sparse $(\alpha, k)$-tuple in $G$.

**Lemma 3.5.9.** Let $G$ be a graph, and let $H$ be a graph. Let $S \subseteq H$ be a stable set, $|S| = k$; and let $m, \delta, c, s > 0$ such that $G$ is $(\delta c^s, H)$-saturated. Then either $G$ contains a $c^m$-sparse $(\delta^2 \cdot k^k c^s, k)$-tuple, or there is an $S$-successor $H'$ of $H$ such that $G$ is $(\frac{1}{2} \delta c^s + m, H')$-saturated.

**Proof.** Let $l = |V(H)|$. For a copy $(T, \phi)$ of $H$, we say that $(\phi(H \setminus S), \phi|_{H \setminus S})$ is the anchor of $(T, \phi)$. Let $T$ denote the set of all anchors of copies of $H$ in $G$. It follows that $|T| \leq ln^{1-k}$.

Let $(S', \phi') \in T$. We say that $(S', \phi')$ is rich if there are at least $\frac{1}{2} \delta c^s n^k$ distinct copies of $H$ in $G$ with anchor $(S', \phi')$. Since $|T| \leq n^{1-k}$, it follows that there are at most $\frac{1}{2} \delta c^s n^l$ distinct copies of $H$ whose anchors are not rich. Consequently, there are at least $\frac{1}{2} \delta c^s n^l$ copies of $H$ whose anchors are rich.

Let $T'$ be the set of rich anchors of copies of $H$, and let $(S', \phi') \in T'$ be rich. We let $S = \{s_1, \ldots, s_k\}$. Let $S(S', \phi')$ be the set of copies of $H$ with anchor $(S', \phi')$. It follows that $\sum_{(S', \phi') \in T'} |S(S', \phi')| \geq \frac{1}{2} \delta c^s n^l$.

For $i \in \{1, \ldots, k\}$, we let $U_i = \{\phi(s_i) : (T, \phi) \in S(S', \phi')\}$. Now let $V_1, \ldots, V_k$ be a random partition of $U = U_1 \cup \cdots \cup U_k$ in which every vertex of $U$ is in $V_i$ with probability $1/k$ independently for all $i \in \{1, \ldots, k\}$. Let $(T, \phi) \in S(S', \phi')$. It follows that the probability that $\phi(s_i) \in V_i$ for all $i \in \{1, \ldots, k\}$ is $1/k^k$. Therefore there is a choice of $V_1, \ldots, V_k$ such that

$$|\{(T, \phi) \in S(S', \phi') : \phi(s_i) \in V_i \forall i \in \{1, \ldots, k\}\}| \geq |S(S', \phi')|/k^k.$$
Fix such a choice of $V_1, \ldots, V_k$. Let $W_i = U_i \cap V_i$ for all $i \in \{1, \ldots, k\}$. It follows that

$$|\{(T, \varphi) \in S(S', \varphi') : \varphi(s_i) \in W_i \forall i \in \{1, \ldots, k\}\}| \geq |S(S', \varphi')|/k^k,$$

since $\varphi(s_i) \in U_i$ for all $(T, \varphi) \in S(S', \varphi')$ and $i \in \{1, \ldots, k\}$.

Since $|\{(T, \varphi) \in S(S', \varphi') : \varphi(s_i) \in W_i \forall i \in \{1, \ldots, k\}\}| \leq \prod_{i=1}^k |W_i|$, and since $|S(S', \varphi')|/k^k \geq \frac{1}{2^k \delta c^s n^k}$, it follows that $|W_i| \geq \frac{1}{2^k \delta c^s n}$ for all $i \in \{1, \ldots, k\}$. If for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $W_i, W_j$ is a $c^m$-sparse pair, then $W_1, \ldots, W_k$ is a $c^m$-sparse $(\frac{\delta}{2^k \delta c^s + m}, k)$-tuple, and the statement of the lemma follows. Therefore, we may assume that there exist distinct $i, j \in \{1, \ldots, k\}$ such that $W_i, W_j$ is not $c^m$-sparse. For $w = (w_1, \ldots, w_k)$ with $w_i \in W_i$ for all $i' \in \{1, \ldots, k\}$, we let $X(w, S', \varphi') = (S' \cup \{w_1, \ldots, w_k\}, \varphi_w)$, where $\varphi_w(v) = \varphi'(v)$ if $v \in V(H) \setminus S$, and $\varphi_w(s_i) = w_i$ for all $i' \in \{1, \ldots, k\}$. Let

$$X(S', \varphi') = \{X(w, S', \varphi') : w = (w_1, \ldots, w_k) \in \Pi_{i'1, \ldots, k} W_i, w_i w_j \in E(G)\}.$$

It follows that $|X(S', \varphi')| \geq c^m \Pi_{i=1}^k |W_i|$, since $W_i, W_j$ is not $c^m$-sparse. By definition, every $X(w, S', \varphi') \in X$ is a copy of an $S$-successor of $H$. It follows that $c^m |S(S', \varphi')|/k^k \leq c^m \Pi_{i=1}^k |W_i| \leq |X(S', \varphi')|.$

Since $\sum_{(S', \varphi') \in T} |S(S', \varphi')| \geq \frac{1}{2} \delta c^s n^l$, it follows that $\sum_{(S', \varphi') \in T} |X(S', \varphi')| \geq \frac{1}{2^k \delta c^s + m n^l}$. This implies that $G$ contains at least $\frac{1}{2^k \delta c^s + m n^l}$ distinct copies of $H$-successors. By Lemma 3.5.8 it follows that $H$ has at most $2^{k^2}$ distinct $S$-successors, and therefore, there is an $S$-successor $H'$ of $H$ such that $G$ contains at least $\frac{1}{2^{2^k \delta c^s + m n^l}$ distinct copies of $H'$. It follows that $G$ is $(\frac{1}{2^{2^k \delta c^s + m}, H'})$-saturated. This concludes the proof.

\[\square\]
Proof. We prove this by induction on $L$. For $L = 1$, $H$ consists of a single vertex, and so $G$ is $(1, H)$-saturated. Now let $L > 1$, and let $v \in V(H)$. By induction, it follows that $G$ is $(2^{-L+2}, H \setminus \{v\})$-saturated. Let $(T, \varphi)$ be a copy of $H \setminus \{v\}$ in $G$. Since $G$ is $\varepsilon$-sparse, it follows that $|V(G) \setminus N[T]| \geq |V(G)|/2$. It follows that there are at least $|V(G)|/2$ vertices $w$ such that the function $\varphi_w$, defined as $\varphi_w(x) = \varphi(x)$ if $x \in V(H) \setminus \{v\}$, and $\varphi(v) = w$, yields a copy $(T \cup \{w\}, \varphi_w)$ of $H$. Therefore, there are at least $2^{-L+1}|V(G)||V(H)|$ copies of $H$ in $G$, and so $G$ is $(2^{-L+1}, H)$-saturated.

Lemma 3.5.11. Let $k \in \mathbb{N}$ with $k \geq 2$, and let $H$ be a graph. Then there exist $s \in \mathbb{N}$ and $\varepsilon, \delta > 0$ such that for every $\varepsilon$-sparse graph $G$, and for all $c > 0$ with $c \leq 1/2$, either $G$ is $(\delta c^s, H)$-saturated, or $G$ contains a $c$-sparse $(\delta c^s, k)$-tuple.

Proof. Let $L \in \mathbb{N}$ such that $H$ is $(L, k)$-forcible. Let $s = \left(\frac{L}{2}\right) + 1$. Let $\varepsilon = 1/(2L)$ and let

$$\delta = \left(\frac{1}{k^2 2^{2^2 + 1}}\right)^s 2^{-L+1}.$$

Let $G$ be $\varepsilon$-sparse.

We now play the $k$-tuple game on $L$ vertices. Since every step adds an edge, it follows that there are at most $s$ rounds. For $i \in \{0, \ldots, s\}$, we let $\delta_i = \left(\frac{1}{k^2 2^{2^2 + 1}}\right)^i 2^{-L+1}$. Then $\delta_{i+1} \leq \frac{1}{2^{k^2 2^{2^2}} \delta_i}$. Let $H_0$ be an $L$-vertex graph with no edges. By Lemma 3.5.10, it follows that $G$ is $(\delta_0, H_0)$-saturated.

For $i \in \{1, \ldots, s\}$, we assume by induction that $H_{i-1}$ is the graph in round $i$ of the $k$-tuple game on $L$ vertices in which player $A$ is using a strategy that forces $H$, and $G$ is $(\delta_{i-1} c^{i-1}, H_{i-1})$-saturated. If player $A$ stops the game in this round, then $H_{i-1}$ contains $H$ as an induced subgraph, and $G$ is $(\delta c^s, H_{i-1})$-saturated. Since $H$ is an induced subgraph of $H$, it follows that $G$ is $(\delta c^s, H)$-saturated by Lemma 3.5.7 and the result of the lemma follows. Therefore, we may assume that player $A$ selects a stable set $S \subseteq V(H_{i-1})$ of size $k$. We now apply Lemma 3.5.9 with $\delta_{i-1}, c, i - 1$ and $m = 1$ to $H_{i-1}$. It follows that either $G$ has a $c$-sparse $(\delta_i c^{i-1}, k)$-tuple, and we
are done, or there is an $S$-successor $H'$ of $H_{i-1}$ such that $G$ is $(\delta_c^i, H')$-saturated. We let $H'$ be the graph that player $B$ selects as $H_i$, and continue the $k$-tuple game. Since the $k$-tuple game terminates in at most $s$ rounds, and since player $A$ is using a strategy that forces $H$, it follows that when this $k$-tuple game terminates, either $G$ is $(\delta c^s, H)$-saturated, or $G$ has a $c$-sparse $(\delta c^s, k)$-tuple. This concludes the proof.

**Lemma 3.5.12.** Let $k \in \mathbb{N}$ with $k \geq 2$, and let $H$ be a graph with $s$ edges. Then there exist $\varepsilon, \delta > 0$ such that for every $\varepsilon$-sparse graph $G$, and for all $c > 0$ with $c \leq 1/2$, either $G$ is $(\delta c^s, H)$-saturated, or $G$ contains a $c$-sparse $(\delta c^s, 2)$-tuple.

**Proof.** This follows from Lemma 3.5.11 since $H$ is $(l, 2)$-forcible and player $A$ has a strategy that forces $H$ in $s + 1$ rounds.

The proof of the following lemma is similar to the proof of Lemma 3.5.11 but yields a slightly different result.

**Lemma 3.5.13.** Let $k, t \in \mathbb{N}$ and let $H$ be an $l$-vertex graph. Then there exist $S, \varepsilon, \delta > 0$ such that for every $\varepsilon$-sparse graph $G$, and for all $c > 0$ with $c \leq 1/2$, either $G$ is $(\delta c^s, H)$-saturated, or $G$ contains a $c^{s+t}$-sparse $(\delta c^s, k)$-tuple for some $s \in \{0, \ldots, S\}$.

**Proof.** Let $L \in \mathbb{N}$ such that $H$ is $(L, k)$-forcible. Let $s = \binom{L}{2} + 1$, and let $T \in \mathbb{N}$ such that $2^T \geq t$. Let $S = 2^{s+T}$. Let $\varepsilon = 1/(2L)$ and let

$$\delta = \left(\frac{1}{k^2 k^{2^2 + 1}}\right)^s \cdot 2^{-L+1}.$$ 

Let $G$ be $\varepsilon$-sparse, and let $c > 0$. We may assume that $c \leq 1$.

We now play the $k$-tuple game on $L$ vertices. Since every step adds an edge, it follows that there are at most $s$ rounds. For $i \in \{0, \ldots, s\}$, we let $\delta_i = \left(\frac{1}{k^i k^{2^i + 1}}\right)^s \cdot 2^{-L+1}$ and $s_i = 2^{i+T} - t$. Then $\delta_{i+1} \leq \frac{1}{2 k^i k^{2^i}} \delta_i$ if $i < s$, and $s_i \in \{0, \ldots, S\}$ for all
\(i \in \{0, \ldots, s\}\). Let \(H_0\) be an \(L\)-vertex graph with no edges. By Lemma 3.5.10, it follows that \(G\) is \((\delta_0 c^{s_0}, H_0)\)-saturated.

For \(i = \{1, \ldots, s\}\), we assume by induction that \(H_{i-1}\) is the graph in round \(i\) of the \(k\)-tuple game on \(L\) vertices in which player \(A\) is using a strategy that forces \(H\), and \(G\) is \((\delta_i c^{s_i}, H_{i-1})\)-saturated. If player \(A\) stops the game in this round, then \(H_{i-1}\) contains \(H\) as an induced subgraph, and \(G\) is \((\delta c^S, H_{i-1})\)-saturated. Since \(H\) is an induced subgraph of \(H_{i-1}\), it follows that \(G\) is \((\delta c^S, H)\)-saturated by Lemma 3.5.7, and the result of the lemma follows.

Therefore, we may assume that player \(A\) selects a stable set \(T \subseteq V(H_{i-1})\) of size \(k\). We now apply Lemma 3.5.9 with \(m = s_{i-1} + t, \delta = \delta_{i-1}, c,\) and \(s = s_{i-1}\) to \(H_{i-1}\). It follows that either \(G\) has a \(c^{s_{i-1}+t}\)-sparse \((\delta_i c^{s_{i-1}}, k)\)-tuple, and we are done, or, since \(s_i \geq s_{i-1} + s_{i-1} + t\), there is a \(T\)-successor \(H'\) of \(H_{i-1}\) such that \(G\) is \((\delta_i c^{s_i}, H')\)-saturated. We now let \(H'\) be the graph that player \(B\) selects as \(H_i\), and continue the \(k\)-tuple game. Since the \(k\)-tuple game terminates in at most \(s\) rounds, and since player \(A\) is using a strategy that forces \(H\), it follows that when this \(k\)-tuple game terminates, either \(G\) is \((\delta c^S, H)\)-saturated, or \(G\) has a \(c^{s_i+t}\)-sparse \((\delta c^{s_i}, k)\)-tuple for some \(i \in \{0, \ldots, s\}\). This concludes the proof.

The following lemma proves that we can get \(k\) sets, each of size \(n^\varepsilon\), pairwise anticomplete, if we exclude a graph \(H\) as an induced subgraph of an \(\varepsilon\)-sparse graph. This is similar to Theorem 3.4.1, except that we assume sparsity and guarantee an anticomplete pair.

**Lemma 3.5.14.** Let \(H\) be a graph and \(k \in \mathbb{N}\). Then there exists an \(\varepsilon > 0\) such that every \(\varepsilon\)-sparse \(H\)-free graph \(G\) with \(|V(G)| = n\) contains a \(0\)-sparse \((\varepsilon n^{\varepsilon-1}, k)\)-tuple.

**Proof.** Let \(\varepsilon', \delta, s > 0\) be as in Lemma 3.5.11. Let \(c = \frac{1}{2^k+1} n^{-1/(s+1)}\). Let

\[
\varepsilon = \min \left\{ \varepsilon', \frac{1}{2k} \cdot \frac{1}{s+1}, \frac{\delta}{k} \cdot \frac{1}{2^k}, \frac{1}{2^k+1} \cdot \frac{1}{k^2} \right\},
\]

\[153\]
and let $G$ be $\epsilon$-sparse and $H$-free. By Lemma 3.5.10, it follows that $G$ contains a stable set of size $k$; therefore, we may assume that $\epsilon n^\epsilon > 1$.

By Lemma 3.5.11, it follows that either $G$ is $(\delta c^\delta, H)$-saturated or $G$ contains a $c$-sparse $(\delta c^\delta, k)$-tuple. Since $G$ is $H$-free, it follows that $G$ is not $(\delta c^\delta, H)$-saturated, and therefore $G$ contains a $c$-sparse $(\delta c^\delta, k)$-tuple $A_1, \ldots, A_k$.

For $i, j \in \{1, \ldots, k\}$ with $i \neq j$, we let $B_{ij}$ be the set of vertices $v \in A_i$ such that $j$ has at least $c|A_j|/k$ neighbors in $A_j$. Since

$$ck|B_{ij}||A_j| \leq |E(B_{ij}, A_j)| \leq |E(A_i, A_j)| \leq c|A_i||A_j|,$$

it follows that $|B_{ij}| \leq |A_i|/k$. For $i \in \{1, \ldots, k\}$, we let $C_i = A_i \setminus \left( \bigcup_{j \in \{1, \ldots, k\} \setminus \{i\}} B_{ij} \right)$.

It follows that $|C_i| \geq \delta c^\delta n/k$ for all $i \in \{1, \ldots, k\}$. It follows that $|C_i| \geq 2k^k \epsilon n^\epsilon$ for all $i \in \{1, \ldots, k\}$. Moreover, for all $v \in C_i$, we have $|N(v) \cap C_j| \leq k|A_j| \leq k^2 c|C_j| \leq k^2 \frac{1}{2^k+1} n^\epsilon|C_j|$ for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$.

Let $k \in \mathbb{N}$, and let $\alpha, \epsilon > 0$. Let $C_1, \ldots, C_k$ be pairwise disjoint, with $|C_i| \geq 2k^k \epsilon n^\epsilon$ for all $i \in \{1, \ldots, k\}$, and such that $|N(v) \cap C_j| \leq n^{-k} |C_j|$ for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, and for all $v \in C_i$, and suppose that $\alpha \leq \frac{1}{2^k+1} \epsilon$ and $\epsilon n^\epsilon \geq 1$. Then there exist pairwise anticomplete sets $D_1, \ldots, D_k$ with $D_i \subseteq C_i$ and $|D_i| \geq \epsilon n^\epsilon$ for all $i \in \{1, \ldots, k\}$.

We prove this by induction on $k$. For $k = 1$, it is true by letting $C_1 = D_1$. Now let $k > 1$, and let $D_1 \subseteq C_1$ be such that $|D_1| = \lfloor \epsilon n^\epsilon \rfloor \leq 2\epsilon n^\epsilon$. It follows that

$$|N(D_1) \cap C_i| \leq \alpha n^{-k} k^2 |D_1||C_i| \leq 2k^2 \epsilon |C_j| \leq |C_j|/2.$$

We apply induction, with $\alpha' = 2\alpha$, to $C_2 \setminus N(D_1), \ldots, C_k \setminus N(D_1)$. Since $|C_i| \geq 2k^k \epsilon n^\epsilon$, it follows that

$$|C_i \setminus N(D_1)| \geq |C_i|/2 \geq 2^{k-1} \epsilon n^\epsilon.$$
for all \( i \in \{1, \ldots, k\} \), and \( \alpha' \leq \frac{1}{2(k-1)+1} \). Since \(|C_j \setminus N(D_1)| \geq |C_j|/2\) for all \( j \in \{2, \ldots, k\} \), it follows that \(|N(v) \cap (C_j \setminus N(D_1))| \leq \alpha n^{-\epsilon} k^2 |C_j| \leq \alpha' n^{-\epsilon} k^2 |C_j \setminus N(D_1)|\) for all \( v \in C_i \) and \( i, j \in \{2, \ldots, k\} \) with \( i \neq j \). It follows that there exist sets \( D_2, \ldots, D_k \), pairwise anticomplete, with \( D_i \subseteq C_i \) and \( |D_i| \geq \epsilon n^\varepsilon \) for all \( i \in \{2, \ldots, k\} \). But then \( D_1, \ldots, D_k \) are the desired sets. This proves (3.14).

Now we let \( \alpha = \frac{1}{2^k+1} \); then the statement of the lemma follows from (3.14).

The next lemma is an improvement of Corollary 3.4.4.

**Lemma 3.5.15.** Let \( H \) be a graph. Then there exist \( \varepsilon, \delta > 0 \) such that if \( G \) is \( H \)-free and \( \varepsilon \)-sparse, then \( G \) has an anticomplete pair \( A, B \) with \(|A||B| \geq \delta n^{1+\varepsilon}\).

**Proof.** Let \( S, \delta, \varepsilon \) be as in Lemma 3.5.13 with \( k = 2 \) and \( t = 1 \). We may assume that \( \varepsilon \leq 1/(2S+1) \); we let \( \delta' = \min\{\varepsilon, \delta/8\} \). We claim that this \( \varepsilon \) and \( \delta'^2 \) satisfy the conditions of the lemma.

Let \( G \) be \( \varepsilon \)-sparse, \(|V(G)| = n\), and let \( c = n^{-1/(2S+1)} \). It follows from Lemma 3.5.13 that either \( G \) is \((\delta c^S, H)\)-saturated, or \( G \) contains a \( c^{s+1}\)-sparse \((\delta c^s, 2)\)-tuple for some \( s \in \{0, \ldots, S\} \). Since \( G \) is \( H \)-free, it follows that \( G \) contains a \( c^{s+1}\)-sparse \((\delta c^s n, \delta c^s n)\)-pair \( A, B \) for some \( s \in \{0, \ldots, S\} \).

Let \( A' \subseteq A \) be the set of vertices in \( A \) with at least \( 2c^{s+1}|B| \) neighbors in \( B \). It follows that since \(|E(A, B)| \leq c^{s+1}|A||B|\), we have \(|A'| \leq |A|/2\), and so \(|A \setminus A'| \geq \frac{\delta}{2} c^s n \geq \frac{\delta}{2} n^{1-s/(2S+1)} \geq \frac{\delta}{2} n^{(s+1)/(2S+1)} \geq \frac{\delta}{2} n^{(s+1)\varepsilon} \).

We may assume that \( \delta'^n^{\varepsilon} > \frac{1}{2} \), for otherwise, \( \{v\}, V(G) \setminus N[v] \) is an anticomplete pair and \(|\{v\}| |V(G) \setminus N[v]| \geq (1 - \varepsilon)n2\delta'n^{\varepsilon} \geq \delta'n^{1+\varepsilon} \), and the statement of the lemma holds.
Now let $A^* \subseteq A \setminus A'$ with $|A^*| = \lceil \delta' n^{(s+1)/(2S+1)} \rceil$. It follows that since $\delta' n^{(s+1)/(2S+1)} \geq \delta n^\epsilon$, we have $|A^*| \leq 2\delta' n^{(s+1)/(2S+1)}$. Therefore,

$$|N(A^*) \cap B| \leq 2c^{s+1}|A^*||B| \leq 4\delta' c^{s+1} n^{(s+1)/(2S+1)}|B| = 4\delta'|B| = \frac{\delta}{2}|B|.$$ 

Let $B^* = B \setminus N(A^*)$. It follows that

$$|B^*| \geq \left(1 - \frac{\delta}{2}\right)|B| \geq |B|/2 \geq \delta' c^s \geq \delta n^{1-s/(2S+1)},$$

and so $A^*, B^*$ is an anticomplete pair with

$$|A^*||B^*| \geq \delta'^2 n^{(s+1)/(2S+1)+(1-s/(2S+1))} = \delta'^2 n^{1+1/(2S+1)} \geq \delta'^2 n^{1+\epsilon}.$$ 

3.5.3 Bipartite graphs and matchings

In this section, we prove that a strengthening of Question 3.3.2 is true for all bipartite graphs $H$.

**Lemma 3.5.16.** Let $H$ be bipartite. Then there exist $t, s \in \mathbb{N}$ and $\delta, \epsilon > 0$ such that for all $c > 0$ with $c \leq 1/2$, if $G$ is $\epsilon$-sparse and not $(\delta c^s, H)$-saturated, then $G$ contains a $c$-sparse $(\epsilon' c^s n, \epsilon n)$-pair.

**Proof.** Let $A, B$ be a bipartition of $H$ with $|A| = a, |B| = b$. Let $S, \epsilon, \delta$ be as in Lemma 3.5.13 with $t = b + 1, k = a$. We may assume that $\epsilon \leq 1/(2^{b+3} \cdot a^4)$. Let $\delta' = \delta/(k^4 \cdot 2^{b+3})$ and let $\epsilon' = \epsilon \cdot \delta'$. We claim that for all $c > 0$, if $G$ is $\epsilon'$-sparse, then either $G$ is $(\epsilon' c^{(S+b)a}, H)$-saturated, or $G$ contains a $c$-sparse $(\epsilon' c^{S+b} n, \epsilon' n)$-pair.

Let $G$ be $\epsilon'$-sparse and $c > 0$. By Lemma 3.3.7 we may assume that $c \leq 2\epsilon$. By Lemma 3.5.13 it follows that either $G$ is $(\delta c^S, H)$-saturated, or $G$ contains a
\(c^{s+b+1}\)-sparse \((\delta c^s, a)\)-tuple for some \(s \in \{0, \ldots, S\}\). If the former, then the statement of the lemma follows. Therefore, we may assume that \(G\) contains a \(c^{s+b+1}\)-sparse \((\delta c^s, a)\)-tuple \(A_1, \ldots, A_a\).

For \(i, j \in \{1, \ldots, k\}\) with \(i \neq j\), let \(B_{ij}\) be the set of vertices in \(A_i\) with at least \(kc^{s+b+1}|A_j|\) neighbors in \(A_j\). Since

\[
k_c^{s+b+1}|A_j||B_{ij}| \leq |E(A_i, A_j)| \leq c^{s+b+1}|A_i||A_j|,\]

it follows that \(|B_{ij}| \leq |A_i|/k\) for all \(i \in \{1, \ldots, k\}\). For \(i \in \{1, \ldots, k\}\), we let \(C_i = A_i \setminus (\bigcup_{j \in \{1, \ldots, k\} \setminus \{i\}} B_{ij})\). It follows that \(|C_i| \geq |A_i|/k\) for all \(i \in \{1, \ldots, k\}\), and consequently, \(|N(v) \cap C_j| \leq k^2c^{s+b+1}|C_j|\) for all \(v \in C_i\) and all \(i, j \in \{1, \ldots, k\}\) with \(i \neq j\).

Let \(H\) be a bipartite graph with bipartition \(A, B\) with \(|B| = b\). Let \(A = \{a_1, \ldots, a_k\}\), let \(s \in \mathbb{N}, \varepsilon, \delta, c > 0\) be such that \(\varepsilon \leq \delta\) and \(G\) is \(\varepsilon\)-sparse and \(c \leq 2\varepsilon \leq 1/(2^{b+2}k^4)\), and let \(A_1, \ldots, A_k\) be pairwise disjoint with \(|A_i| \geq \delta c^s n\) for all \(i \in \{1, \ldots, k\}\). Suppose further that \(|N(v) \cap A_j| \leq k^2c^{s+b+1}|A_j|\) for all \(v \in A_i\) and all \(i, j \in \{1, \ldots, k\}\) with \(i \neq j\). Then (3.15)

157

either

- there are at least \(\frac{\delta^k c^{(s+b)k} n^{|V(H)|}}{2^{2b+1}}\) distinct copies \((T, \varphi)\) of \(H\) in \(G\) with \(\varphi(a_i) \in A_i\) for all \(i \in \{1, \ldots, k\}\), and \(\varphi(b') \in V(G) \setminus (\bigcup_{i \in \{1, \ldots, k\}} A_i)\) for all \(b' \in B\); or

- \(G\) has a \(c\)-sparse \((\varepsilon c^{s+b} n, \varepsilon n)\)-pair.

We prove this by induction on \(|B|\). Suppose that \(B = \emptyset\). Let \((c_1, \ldots, c_k) \in \Pi_{i \in \{1, \ldots, k\}} A_i\) be chosen at random. For \(i, j \in \{1, \ldots, k\}\), the probability that \(c_i c_j \in E(G)\) is at most \(k^2c^{s+1}\). It follows that the probability that \(\{c_1, \ldots, c_k\}\) is stable is
at least $1 - k^4c^{s+1} \geq 1/2$. It follows that at least half of the $k$-tuples $(c_1, \ldots, c_k) \in \Pi_{i \in \{1, \ldots, k\}} A_i$ yield copies of $(T, \varphi)$ of $H$ with $\varphi(a_i) \in A_i$ for all $i \in \{1, \ldots, k\}$. We have found $\frac{\delta k}{2} c^s n^k$ such copies of $H$, and hence (3.15) holds when $B = \emptyset$.

Now let $|B| > 0$, and let $v \in B$. Let $I \subseteq \{1, \ldots, k\}$ be such that $i \in I$ if and only if $v$ is adjacent to $a_i$. For $i \in \{1, \ldots, k\}$, we define $D_i$ as follows:

- if $i \in I$, we let $D_i$ be the set of vertices in $V(G) \setminus A_i$ with at most $c|A_i|$ neighbors in $A_i$;
- if $i \not\in I$, we let $D_i$ be the set of vertices in $V(G) \setminus A_i$ with at least $4k\varepsilon|A_i|$ neighbors in $A_i$.

It follows that $|D_i| \leq \varepsilon n$ if $i \in I$, for otherwise $A_i, D_i$ is a $c$-sparse $(\varepsilon c^* n, \varepsilon n)$-pair, and (3.15) follows. If $i \in I$, then $|D_i| \leq |V(G)|/(4k)$, since

$$4k\varepsilon|A_i||D_i| \leq |E(A_i, V(G) \setminus N[A_i])| \leq \varepsilon|V(G)||A_i|.$$ 

Since $\varepsilon \leq 1/(4k)$, it follows that $|A_i| \leq |V(G)|/(4k)$ for all $i \in \{1, \ldots, k\}$, and consequently, $C = V(G) \setminus (\bigcup_{i \in \{1, \ldots, k\}} (D_i \cup A_i))$ satisfies $|C| \geq |V(G)|/2$.

Now let $d \in C$, and for $i \in \{1, \ldots, k\}$, we let $A'_i$ be defined as follows:

- if $i \in I$, we let $A'_i = N(d) \cap A_i$;
- if $i \not\in I$, we let $A'_i = A_i \setminus N(d)$.

Since $c \leq 2\varepsilon \leq 1/(4k)$, it follows that $|A'_i| \geq c|A_i|$ for all $i \in \{1, \ldots, k\}$ by the choice of $C$.

We apply induction to

$$H \setminus v, A'_1, \ldots, A'_k, G \setminus \left( N[d] \setminus \left( \bigcup_{i \in \{1, \ldots, k\}} A_i \right) \right), \varepsilon' = 2\varepsilon, \delta, c, s + 1;$$ 

158
the choice of $\varepsilon'$ ensures that $G \setminus N[d]$ is $\varepsilon'$-sparse. This is valid since $|N(v) \cap A'_j| \leq k^2 \varepsilon^s + b + 1 |A_j| \leq k^2 \varepsilon^s |A'_j|$.

It follows that either $G$ has a $c$-sparse $(\varepsilon' c^{s+1+b-1}, \varepsilon' n)$-pair (in which case (3.15) follows since $\varepsilon' \geq \varepsilon$), or there are at least $\delta k^2 b c^{s k} n |V(H)|^{-1}$ distinct copies $(T, \varphi)$ of $H \setminus \{v\}$ in $G \setminus N[v]$ with $\varphi(a_i) \in A_i$ for all $i \in \{1, \ldots, k\}$, and $\varphi(b') \in V(G) \setminus (N[d] \cup \bigcup_{i \in \{1, \ldots, k\}} A'_i)$ for all $b' \in B$.

For each such copy $(T, \varphi)$, we get a copy of $H$ by letting $\varphi(v) = d$, since $d$ is non-adjacent to $\varphi(b')$ for all $b' \in B$, and for all $i \in \{1, \ldots, k\}$, $\varphi(v)$ is adjacent to $a_i$ if and only if $i \in I$. This proves that for every $d \in C$, $G$ contains $\delta k^2 b c^{s k} n |V(H)|^{-1}$ distinct copies $(T, \varphi)$ of $H$ in $G \setminus N[v]$ with $\varphi(a_i) \in A_i$ for all $i \in \{1, \ldots, k\}$, and $\varphi(b') \in V(G) \setminus \left( \bigcup_{i \in \{1, \ldots, k\}} A'_i \right)$ for all $b' \in B$.

Since $|C| \geq |V(G)|/2$, it follows that $G$ contains $\delta k^2 b c^{s k} n |V(H)|^{-1}$ distinct copies $(T, \varphi)$ of $H$ in $G \setminus N[v]$ with $\varphi(a_i) \in A_i$ for all $i \in \{1, \ldots, k\}$, and $\varphi(b') \in V(G) \setminus \left( \bigcup_{i \in \{1, \ldots, k\}} A'_i \right)$ for all $b' \in B$. This proves (3.15).

Now the result of the Lemma follows by applying (3.15) to $C_1, \ldots, C_k$ with $\delta' = \delta/(k^4 \cdot 2^{b+3})$ and $\varepsilon' = \varepsilon \cdot \delta'$.

Let $G$ be a graph, and let $C, D \subseteq V(G)$ be disjoint. A partition $(S, R)$ of a set $D$ is $(C, j)$-split for $j \in \mathbb{N}$ if

- $|N(v) \cap C| \leq j$ for all $v \in S$;
- $|N(v) \cap C| \geq j$ for all $v \in R$; and
- $|S|, |R| \geq |D|/4$.

**Lemma 3.5.17.** Let $G$ be a graph, $k \in \mathbb{N}$, and $D \subseteq V(G)$ with $|D| \geq 4^k + 1$. Let $C_1, \ldots, C_k \subseteq V(G)$ be a collection of sets, and let $D \subseteq V(G)$. Then there is a set $I \subseteq \{1, \ldots, k\}$ with $|I| \geq k/2$, and a set $D' \subseteq D$ and $j_i \in \mathbb{N}$ for all $i \in I$, and
a partition \((A, B)\) of \(D'\) such that \(|A|, |B| \geq |D|/4^{k+1}\) and for all \(i \in I\), \((A, B)\) is \((C_i, j_i)\)-split.

**Proof.** We start by proving the following claim.

\[\text{Let } G \text{ be a graph, } k \in \mathbb{N}, \text{ and } D \subseteq V(G) \text{ with } |D| \geq 4^{k+1}. \text{ Let } C_1, \ldots, C_k \subseteq V(G) \text{ be a collection of sets, and let } D \subseteq V(G). \text{ Then}
\]

\[\text{there is a set } D' \subseteq D \text{ and } j_1, \ldots, j_k \in \mathbb{N}, \text{ and a partition } (A, B) \text{ of } D' \text{ such that } |A|, |B| \geq |D|/4^{k+1} \text{ and for all } i \in \{1, \ldots, k\}, \text{ either } (A, B) \text{ or } (B, A) \text{ is } (C_i, j_i)\)-split.\]

We prove this by induction on \(k\). For \(k = 0\), we let \(A, B\) be a partition of \(D\) with \(|A|, |B| \in \{\lceil |D|/2 \rceil, \lfloor |D|/2 \rfloor\}\). Since \(|D| \geq 4\), it follows that \(|A|, |B| \geq |D|/4\).

Now let \(k \geq 1\). We apply induction to \(C_1, \ldots, C_{k-1}, D\). It follows that there is a set \(D' \subseteq D\) and a partition \((A, B)\) of \(D'\), and \(j_1, \ldots, j_{k-1} \in \mathbb{N}\) such that \(|A|, |B| \geq |D|/4^k\) and for all \(i \in \{1, \ldots, k-1\}\), either \((A, B)\) or \((B, A)\) is \((C_i, j_i)\)-split. By deleting vertices from \(A, B\), we may assume that \(|A| = |B|\). Let \(j_k\) be the median of \((|N(v) \cap C_k|)_{v \in D'}\). Then there exist \(A', B' \subseteq D'\) with \(|A'|, |B'| \in \{\lceil |D'|/2 \rceil, \lfloor |D'|/2 \rfloor\}\) such that \((A', B')\) is \((C_k, j_k)\)-split. Since \(|D'| \geq 4\), it follows that \(|A'|, |B'| \geq |D'|/4\).

By symmetry, we may assume that \(|A'| \leq |B'|\).

We consider two cases. Suppose first that \(|A' \cap A| \geq |A' \cap B|\). It follows that

\[|A' \cap A| \geq |A'|/2 \geq |D'|/8 \geq \frac{2|D|}{4^k \cdot 8} = |D|/4^{k+1}.\]

Moreover,

\[|B' \cap B| = |B| - |A' \cap B| \geq |B| - \frac{|A'|}{2} \geq |B| - \frac{|D'|}{4} \geq |B| - \frac{|B|}{2} \geq |D|/4^{k+1},\]

and therefore, \((A' \cap A, B' \cap B)\) is the desired partition, and \((3.16)\) holds.

\[\text{160}\]
Now we may assume that \(|A' \cap A| \leq |A' \cap B|\). It follows that
\[
|A' \cap B| \geq |A'|/2 \geq |D'|/8 \geq \frac{2|D|}{4^k} \cdot \frac{1}{4} = |D|/4^{k+1}.
\]
Moreover,
\[
|B' \cap A| = |A| - |A' \cap A| \geq |A| - \frac{|A'|}{2} \geq |A| - \frac{|D'|}{4} \geq |A| - \frac{|A|}{2} \geq |D|/4^{k+1},
\]
and therefore, \((A' \cap B, B' \cap A)\) is the desired partition, and again, (3.16) holds.

Now we apply (3.16) to \(C_1, \ldots, C_k\) and obtain \((A, B), D'\) and \(j_1, \ldots, j_k\) as in (3.16). We let \(I_1\) be the set of \(i \in \{1, \ldots, k\}\) such that \((A, B)\) is \((C_i, j_i)\)-split, and let \(I_2 = \{1, \ldots, k\} \setminus I\). There exists an \(\eta \in \{1, 2\}\) such that \(|I_\eta| \geq k/2\). We let \(I = I_\eta\); then \(I, (A, B), D',\) and \(j_i\) for \(i \in I\) satisfy the statement of the lemma.

Let \(H\) be a graph. We call \(H\) almost bipartite if \(H\) is triangle-free and there is a partition of \(V(H)\) into \(A, B\) such that \(H|A\) is a stable set and \(H|B\) is a graph with maximum degree one. We call such a partition an almost bipartition.

Lemma 3.5.16 can be generalized to almost bipartite graphs, but for the sake of simplicity, we only prove the case of one copy of \(H\).

**Lemma 3.5.18.** Let \(H\) be almost bipartite. Then there exist \(T \in \mathbb{N}\) and \(\varepsilon > 0\) such that for all \(c > 0\) with \(c \leq 1/2\), if \(G\) is \(\varepsilon\)-sparse and \(H\)-free, then \(G\) has a \(c\)-sparse \((\varepsilon c^T n, \varepsilon n)\)-pair.

**Proof.** Let \(A, B\) be a partition of \(V(H)\) with \(|A| = a, |B| = b\) such that \(H|A\) is stable and \(H|B\) has maximum degree one. By enlarging \(H\) if necessary, we may assume that every vertex of \(H|B\) has degree one in \(H|B\). Let \(A = \{a_1, \ldots, a_a\}\) and \(B = \{b_1, \ldots, b_b\}\) such that \(b_{2i-1}\) is adjacent to \(b_{2i}\) for all \(i \in \{1, \ldots, b/2\}\). Let \(R = b(\binom{6a}{a})\). Let \(S, \varepsilon, \delta\) be as in Lemma 3.5.13 with \(t = R + 1, k = 6a\).
We claim that the statement of the lemma is true for $T = S + R + 1$ and

$$
\varepsilon' = \min \left\{ \varepsilon, \frac{\delta}{k}, \frac{1}{2Rk8k^{k+3}}, \frac{1}{4a^2k^2} \right\}.
$$

For a collection of sets $C_1, \ldots, C_k$, and for $I \subseteq \{1, \ldots, k\}$ with $|I| = a$ and $i \in \{1, \ldots, b/2\}$, a pair $(u, v)$ of vertices $u, v \in V(G)$ is an $(I, i)$-bridge if the following statements hold:

- $I = \{i_1, \ldots, i_a\}$ with $i_1 < \cdots < i_a$;
- $uv \in E(G)$;
- for all $j \in \{1, \ldots, a\}$, $u$ is complete to $C_{i_j}$ if $b_{2i-1}$ is adjacent to $a_j$; and $u$ is anticomplete to $C_{i_j}$ otherwise;
- for all $j \in \{1, \ldots, a\}$, $v$ is complete to $C_{i_j}$ if $b_{2i}$ is adjacent to $a_j$; and $v$ is anticomplete to $C_{i_j}$ otherwise; and
- $\{u, v\}$ is anticomplete to $C_i$ for all $i \in \{1, \ldots, k\} \setminus I$.

A pair $(u, v)$ is a bridge if there exist $I \subseteq \{1, \ldots, k\}$ with $|I| = a$ and $i \in \{1, \ldots, b/2\}$ such that $(u, v)$ is an $(I, i)$-bridge.

For a collection $C_1, \ldots, C_k$ of sets, a collection $(u_1, v_1), \ldots, (u_r, v_r)$ of bridges is consistent if there exist $I_1, \ldots, I_r, i_1, \ldots, i_r$ such that the following statements hold:

- for all $j \in \{1, \ldots, r\}$, $(u_j, v_j)$ is an $(I_j, i_j)$-bridge;
- for all $j, j' \in \{1, \ldots, r\}$ with $j \neq j'$, either $I_j \neq I_{j'}$ or $i_j \neq i_{j'}$; and
- for all $j, j' \in \{1, \ldots, r\}$ with $j \neq j'$, $\{u_j, v_j\}$ in anticomplete to $\{u_{j'}, v_{j'}\}$.

Furthermore, we say that the bridges $(u_1, v_1), \ldots, (u_r, v_r)$ complete $I \subseteq \{1, \ldots, k\}$ with $|I| = a$ if additionally, for all $i \in \{1, \ldots, b/2\}$, there exists a distinct $(u_{j_i}, v_{j_i})$ such that $(I_{j_i}, i_{j_i}) = (I, i)$. If the bridges $(u_1, v_1), \ldots, (u_r, v_r)$ are consistent and if
$r \geq R$, then the pigeon-hole principle implies that there is an $I \subseteq \{1, \ldots, k\}$ such that $(u_1, v_1), \ldots, (u_r, v_r)$ complete $I$.

The following statement shows that it suffices to find a consistent collection of bridges such that there exists an $I \subseteq \{1, \ldots, k\}$ completed by this collection of bridges.

Let $C_1, \ldots, C_k$ be a collection of disjoint sets such that there exist $p \in \mathbb{N}$, $k \in \mathbb{N}$, $\delta > 0$, and $c \leq 1/(2 \cdot a^2 k^2)$, and

1. for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $|N(v) \cap C_j| \leq k^2 c |C_j|$ for all $v \in C_i$; and
2. $|C_i| \geq \delta c^p n$;

and let $(u_1, v_1), \ldots, (u_r, v_r)$ be a consistent collection of bridges that completes $I \subseteq \{1, \ldots, k\}$. Then $G$ contains $H$ as an induced subgraph.

Let $I_j, i_j$ for $j \in \{1, \ldots, r\}$ be chosen such that the conditions for a consistent collection of bridges hold, and let $I \subseteq \{1, \ldots, k\}$ such that $(u_1, v_1), \ldots, (u_r, v_r)$ complete $I$. By symmetry, we may assume that $I_1 = \cdots = I_{b/2} = I$. Since the collection of bridges is consistent, it follows that $i_1, \ldots, i_{b/2}$ are pairwise distinct. It follows that $\{i_1, \ldots, i_{b/2}\} = \{1, \ldots, b/2\}$. By symmetry, we may assume that $i_j = j$ for all $j \in \{1, \ldots, b/2\}$; we also may assume that $I = \{1, \ldots, a\}$.

Now let $(c_1, \ldots, c_a) \in \Pi_{i \in \{1, \ldots, a\}} C_i$ be chosen uniformly at random. For $i, j \in \{1, \ldots, a\}$ with $i \neq j$, the probability that $c_i c_j \in E(G)$ is at most $k^2 c$. Therefore, the probability that $\{c_1, \ldots, c_a\}$ is a stable set is at least $1 - k^2 a^2 c \geq 1/2$. It follows that there is a choice of $c_1, \ldots, c_a$ such that $\{c_1, \ldots, c_a\}$ is stable. Now let $\varphi$ be defined as follows: $\varphi(a_i) = c_i$ for all $i \in \{1, \ldots, a\}$, $\varphi(b_{2i-1}) = u_i$ and $\varphi(b_{2i}) = v_i$ for all $i \in \{1, \ldots, b/2\}$. It follows that $(\varphi(V(H)), \varphi)$ is a copy of $H$ in $G$. This proves (3.17).
It remains to show that we can always add a bridge in a consistent way.

Let $C_1, \ldots, C_k$ be a collection of disjoint sets such that there exists an $p \in \mathbb{N}$, $k \in \mathbb{N}$ with $k \geq 6a$, $\delta > 0$, such that $\delta \geq \epsilon$, $\epsilon \leq 1/(2rk \cdot 8^{k+3})$, $c \leq 1/(4 \cdot a^2 k^2)$, and suppose that $G$ is $\epsilon$-sparse, and

\begin{itemize}
  \item for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $|N(v) \cap C_j| \leq k^2 c |C_j|$ for all $v \in C_i$; and
  \item for all $i \in \{1, \ldots, k\}$, $|C_i| \geq \delta \epsilon \rho n$ and $|C_i| \leq \epsilon n$;
\end{itemize}

and that $(u_1, v_1), \ldots, (u_r, v_r)$ is a consistent collection of bridges for $C_1, \ldots, C_k$. Then one of the following holds:

\begin{itemize}
  \item there is an $I \subseteq \{1, \ldots, k\}$ such that $(u_1, v_1), \ldots, (u_r, v_r)$ complete $I$;
  \item there is a $c$-sparse $(\epsilon \rho^{p+1} n, \epsilon n)$-pair in $G$; or
  \item there exist $C'_1, \ldots, C'_k$ and $(u_{r+1}, v_{r+1})$ such that for all $i \in \{1, \ldots, k\}$, $|C'_i| \geq c^2 |C_i|$ and $C'_i \subseteq C_i$, and $(u_1, v_1), \ldots, (u_{r+1}, v_{r+1})$ is a consistent collection of bridges for $C'_1, \ldots, C'_k$.
\end{itemize}

Suppose that there is no $c$-sparse $(\epsilon \rho^{p+1} n, \epsilon n)$ in $G$.

For $i \in \{1, \ldots, k\}$, we define sets $D_i, E_i$ as follows:

\begin{itemize}
  \item we let $D_i$ be the set of vertices in $V(G) \setminus C_i$ with at most $c |C_i|$ neighbors in $C_i$;
  \item we let $E_i$ be the set of vertices in $V(G) \setminus C_i$ with at least $4k \epsilon |C_i|$ neighbors in $C_i$.
\end{itemize}

It follows that $|D_i| \leq \epsilon n$ for all $i \in \{1, \ldots, k\}$, for otherwise $C_i, D_i$ is a $c$-sparse $(\epsilon \rho n, \epsilon n)$-pair, and (3.18) follows. For all $i \in \{1, \ldots, k\}$, it follows that $|E_i| \leq$
\[ |V(G)|/(4k), \text{ since} \]

\[ 4k \varepsilon |C_i| E_i \leq |E(C_i, V(G) \setminus C_i)| \leq \varepsilon |V(G)||C_i|. \]

Since \( \varepsilon \leq 1/(16k) \), it follows that \(|C_i \cup D_i \cup E_i| \leq |V(G)|/(2k)\) for all \(i \in \{1, \ldots, k\}\), and consequently, \(C = V(G) \setminus \left( \bigcup_{i \in \{1, \ldots, k\}} (C_i \cup D_i \cup E_i) \right)\) satisfies \(|C| \geq |V(G)|/2\).

Now let \(D = C \setminus \left( \bigcup_{i \in \{1, \ldots, r\}} (N[u_i] \cup N[v_i]) \right)\). It follows that

\[ |D| \geq |C| - 2r\varepsilon |V(G)| \geq |C| - |V(G)|/4 \geq |V(G)|/4. \]

This implies that \(|D| \geq |V(G)|/4 \geq 1/(4\varepsilon) \geq 4^{k+1}\). By Lemma 3.5.17, it follows that there exists a set \(I \subseteq \{1, \ldots, k\}\) with \(|I| \geq k/2\), and a set \(D' \subseteq D\) and \(j_i \in \mathbb{N}\) for all \(i \in I\), and a partition \((A, B)\) of \(D'\) such that \(|A|, |B| \geq |D|/4^{k+1}\) and for all \(i \in I\), \((B, A)\) is \((C_i, j_i)\)-split.

It follows that \(|A|, |B| \geq |V(G)|/4^{k+2}\). Now let \(uv \in E(G)\) with \(u, v \in D'\). For \(i \in I\), we say that \(u, v\) are \(i\)-incomparable if \(|(N(u) \setminus N(v)) \cap C_i| \geq c^2|C_i|\), and \(|(N(v) \setminus N(u)) \cap C_i| \geq c^2|C_i|\).

\[ \text{There is an edge } uv \in E(G)(A \cup B) \text{ and a set } I' \subseteq I \text{ such that } u, v \text{ are } i\text{-incomparable for all } i \in I', \text{ and } I' \geq |I|/3. \]

Suppose that no such edge exists. We say that \(u \to_i v\) if \(|(N(v) \setminus N(u)) \cap C_i| < c^2|C_i|\). Let \(A_1, A_2\) be a partition of \(A\) with \(|A_1|, |A_2| \geq |A|/2 \geq |A|/4\). For every \(u \in A_1, v \in A_2\) with \(uv \in E(G)\), we let \(I_1 = \{i \in I : u \to_i v\}, I_2 = \{i \in I : v \to_i u\}\), and \(I_3 = I \setminus (I_1 \cup I_2)\). If there is an edge \(uv\) such that \(|I_3| \geq |I|/3\), then \(uv\) is the desired edge. Therefore, we may assume that either \(|I_1| \geq |I|/3\) or \(|I_2| \geq |I|/3\). If the former, we say that \(u \to v\), and if the latter, we say that \(v \to u\).
Since $|A_1|, |A_2| \geq |A|/4 \geq \varepsilon n$, it follows that $A_1, A_2$ is not $c$-sparse. Let $E_1 = \{uv \in E(A_1, A_2) : u \rightarrow v, u \in A_1, v \in A_2\}$ and let $E_2 = E(A_1, A_2) \setminus E_1$. Since $E(A_1, A_2) \geq c|A_1||A_2|$, it follows that there exists an $\eta \in \{1, 2\}$ such that $|E_\eta| \geq \frac{c}{2}|A_1||A_2|$. It follows that there exists a vertex $u \in A_\eta$ such that the set $U = \{v \in A_{3-\eta} : u \rightarrow v\}$ satisfies $|U| \geq \frac{c}{2}|A_{3-\eta}| \geq \varepsilon cn$.

We choose such a vertex $u$ and set $U$. If $U, B \setminus N[U]$ is a $c$-sparse $(\varepsilon cn, \varepsilon n)$ pair, then (3.18) holds; therefore, we may assume that this is not the case, and so $|B \setminus N[U]| \leq \varepsilon n$. It follows that $|B \cap N[U]| \geq |B| - \varepsilon n \geq 2k \varepsilon n$.

For $i \in I$, we let $B_i$ be the set of vertices in $B$ with at most $c|C_i \setminus N(u)|$ neighbors in $C_i \setminus N(u)$. Since $|N(u) \cap C_i| \leq 4k \varepsilon |C_i|$, it follows that

$$|C_i \setminus N(u)| \geq \frac{\delta}{2} \varepsilon n \geq \varepsilon c^{p+1} n$$

for all $i \in I$. Since $C_i \setminus N(u), B_i$ is a $c$-sparse pair for all $i \in I$, it follows that either $C_i \setminus N(u), B_i$ is a $(\varepsilon c^{p+1} n, \varepsilon n)$-pair and (3.18) holds, or $|B_i| \leq \varepsilon n$ for all $i \in I$. We may assume that the latter is the case. Let $B' = B \setminus (\bigcup_{i \in I} B_i)$. It follows that $|B'| \geq |B| - k \varepsilon n \geq k \varepsilon n$.

Now let $w \in B'$, and let $v$ be a neighbor of $w$ in $U$. Let $I' = \{i \in I : u \rightarrow_i v\}$. By the choice of $U$, it follows that $|I'| \geq |I|/3$. Since $w \in B'$, it follows that $w$ has at least $c|C_i \setminus N(u)|$ neighbors in $C_i \setminus N(u)$ for all $i \in I$, and $v$ has at most $c^2|C_i|$ neighbors in $C_i \setminus N(u)$ for all $i \in I'$. Since

$$c|C_i \setminus N(u)| - c^2|C_i| \geq \frac{c}{2} |C_i| - c^2|C_i| \geq \frac{c}{4} |C_i| \geq c^2 |C_i|,$$

for all $i \in I'$, it follows that $v \not\sim w$. Since $(B, A)$ is $(C_i, j_i)$-split for all $i \in I$, and since $v \in A, w \in B$, it follows that $|N(w) \cap C_i| \leq |N(v) \cap C_i|$, and therefore $|(N(v) \setminus N(w)) \cap C_i| \geq |(N(w) \setminus N(v)) \cap C_i| \geq c^2 |C_i|$ for all $i \in I'$. This implies that for all $i \in I'$, $v$ and $w$ are $\varepsilon$-incomparable. This proves our claim (3.19).
We now return to the proof of (3.18). Let \( I_j, i_j \) be such that \((u_j, v_j)\) is an \((I_j, i_j)\)-bridge for all \( j \in \{1, \ldots, k\} \). Let \( uv \in E(G|(A \cup B)) \) and \( I' \) as in (3.19). Since \(|I| \geq k/2\), it follows that \(|I'| \geq |I|/3 \geq k/6 \geq a\), and by deleting numbers from \( I' \), we may assume that \(|I'| = a\). We may assume that \((u_1, v_1), \ldots, (u_r, v_r)\) does not complete \( I' \), for otherwise (3.18) holds. Let \( j^* \in \{1, \ldots, b/2\} \) such that \((I_j, i_j) \neq (I', j^*)\) for all \( j \in \{1, \ldots, r\} \). We let \( I' = \{i_1, \ldots, i_a\} \) with \( i_1 < \cdots < i_a \).

We let

\[
I_0 = (\{1, \ldots, k\} \setminus I') \cup \{i_j : j \in \{1, \ldots, a\}, a_j \notin N(b_{2j-1}) \cup N(b_{2j})\};
\]
\[
I_1 = \{i_j : j \in \{1, \ldots, a\}, a_j \in N(b_{2j-1}) \setminus N(b_{2j})\};
\]
\[
I_2 = \{i_j : j \in \{1, \ldots, a\}, a_j \in N(b_{2j}) \setminus N(b_{2j-1})\}.
\]

Since \( H \) is triangle-free, it follows that \( \{1, \ldots, k\} = I_0 \cup I_1 \cup I_2 \). By definition, we have that \( I_1, I_2 \subseteq I' \).

For \( i \in \{1, \ldots, k\} \), we define \( C'_i \) as follows.

- if \( i \in I_0 \), we let \( C'_i = C_i \setminus (N(u) \cup N(v)) \);
- if \( i \in I_1 \), we let \( C'_i = (N(u) \setminus N(v)) \cap C_i \); and
- if \( i \in I_2 \), we let \( C'_i = (N(v) \setminus N(u)) \cap C_i \).

By the choice of \( I_0, I_1, I_2 \), it follows that \((u, v)\) is an \((I', j^*)\)-bridge for \( C'_1, \ldots, C'_k \). We claim that for all \( i \in \{1, \ldots, k\} \), \( C'_i \geq c^2|C_i| \). Since \( u, v \) are \( i \)-incomparable for all \( i \in I' \), this is true for all \( i \in I' \). Now let \( i \in \{1, \ldots, k\} \setminus I' \). It follows that \( i \in I_0 \).

Since \( u, v \in A \cup B \), it follows that \(|(N(u) \cup N(v)) \cap C_i| \leq 8k\varepsilon|C_i| \), and therefore \(|C'_i| \geq (1 - 8k\varepsilon)|C_i| \geq |C_i|/2 \geq c^2|C_i| \). This proves our claim.

Since \( C'_i \subseteq C_i \) for all \( i \in \{1, \ldots, k\} \), it follows that \((u_1, v_1), \ldots, (u_r, v_r)\) is a consistent collection of bridges for \( C'_1, \ldots, C'_k \). Since \( u, v \in A \cup B \), it follows that \( \{u, v\} \) is anticomplete to \( \bigcup_{i \in \{1, \ldots, r\}} \{u_i, v_i\} \). Moreover, since \((u, v)\) is an \((I', j^*)\)-bridge for
Let $G$ be $\varepsilon'$-sparse and $c > 0$. We may assume that $c \leq 2\varepsilon'$ by Lemma 3.3.7. By Lemma 3.5.13, it follows that either $G$ is $(\delta c^s, H)$-saturated, or $G$ contains a $c^{s+t+1}$-sparse $(\delta c^s, k)$-tuple for some $s \in \{0, \ldots, S\}$. If the former, then the statement of the lemma follows. Therefore, we may assume that $G$ contains a $c^{s+t+1}$-sparse $(\delta c^s, a)$-tuple $A_1, \ldots, A_a$.

For $i, j \in \{1, \ldots, k\}$ with $i \neq j$, let $B_{ij}$ be the set of vertices in $A_i$ with at least $kc^{s+t+1}|A_j|$ neighbors in $A_j$. Since

$$kc^{s+t+1}|A_j||B_{ij}| \leq |E(A_i, A_j)| \leq c^{s+t+1}|A_i||A_j|,$$

it follows that $|B_{ij}| \leq |A_i|/k$ for all $i \in \{1, \ldots, k\}$. For $i \in \{1, \ldots, k\}$, we let $C_i = A_i \setminus \bigcup_{j \in \{1, \ldots, k\}\setminus\{i\}} B_{ij}$. It follows that $|C_i| \geq |A_i|/k$ for all $i \in \{1, \ldots, k\}$, and consequently, $|N(v) \cap C_j| \leq k^2 c^{s+t+1}|C_j|$ for all $v \in C_i$ and all $i, j \in \{1, \ldots, k\}$ with $i \neq j$. Since for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $C_i, C_j$ is a $c^{s+t}$-sparse pair, the result follows if there exists an $i \in \{1, \ldots, k\}$ with $|C_i| \geq \varepsilon n$.

We now apply (3.18) to $C_1, \ldots, C_k$ while the third outcome applies, i.e. we build a sequence \(\{C^0_i\}_{i \in \{1, \ldots, k\}}, \ldots, \{C^r_i\}_{i \in \{1, \ldots, k\}}\) and \((u_1, v_1), \ldots, (u_r, v_r)\) such that

1. for all $i \in \{1, \ldots, r\}$, $C^i_1, \ldots, C^i_k$ arise from applying (3.18) to $C^{i-1}_1, \ldots, C^{i-1}_k$;
2. for all $i \in \{1, \ldots, r\}$ and $j \in \{1, \ldots, k\}$, $|C^i_j| \geq \delta c^{s+2r}|C^0_j|$;
3. for all $i \in \{1, \ldots, r\}$, $(u_1, v_1), \ldots, (u_i, v_i)$ is a consistent collection of bridges for $C^i_1, \ldots, C^i_k$.

We let $C^0_i = C_i$ for $i \in \{1, \ldots, k\}$. Now suppose that we have constructed \(\{C^r_i\}_{i \in \{1, \ldots, k\}}\). By (3.17), it follows that either $r \leq R/2$, or $G$ contains $H$ as an
induced subgraph; we may assume that the former is the case. This implies that for all \( i, j \in \{1, \ldots, k\} \) with \( i \neq j \),

\[ |N(v) \cap C^r_j| \leq k^2 e^{s+t+1}|C_j| \leq k^2 e^{s+2r+1}|C_j| \leq k^2 c|C^r_j| \]

for all \( v \in C^r_i \). Since \( |C^r_j| \geq \delta c^{s+2r}|C^0_j| \geq \delta/kc^{s+2r} \), it follows that we may apply (3.18) with \( p = s + 2r, \delta' = \delta/k, \) and \( \varepsilon' \). If there is an \( I \subseteq \{1, \ldots, k\} \) such that \((u_1, v_1), \ldots, (u_r, v_r)\) complete \( I \), then \( G \) contains \( H \) as an induced subgraph by (3.17).

If there is a \( c \)-sparse \((\varepsilon c^{r+1} n, \varepsilon n)\)-pair in \( G \), then there is a \( c \)-sparse \((\varepsilon c^T n, \varepsilon n)\)-pair in \( G \), and the statement of the lemma follows. Therefore, we may assume that there exist sets \( C'_1, \ldots, C'_k \) and a pair of vertices \((u_{r+1}, v_{r+1})\) such that for all \( i \in \{1, \ldots, k\} \), \( |C'_i| \geq c^2|C_i| \) and \( C'_i \subseteq C_i \), and \((u_1, v_1), \ldots, (u_{r+1}, v_{r+1})\) is a consistent collection of bridges for \( C'_1, \ldots, C'_k \). Now let \( C'^{r+1}_i = C'_i \) for all \( i \in \{1, \ldots, k\} \). Then this collection of sets and bridges satisfies the conditions for \( r + 1 \). Since this process terminates when \( r \geq R/2 \) by (3.17), the result of the lemma follows. \( \square \)

Let \( S \subseteq \mathbb{N} \). A graph \( H \) is an \( S \)-subdivision of a graph \( H' \) if \( H \) arises from \( H' \) by subdividing edges, and each edge of \( H' \) corresponds to a path in \( H \) of length \( k \) for some \( k \in S \).

Lemma 3.5.18 implies the following strengthening of Theorem 3.3.17.

**Lemma 3.5.19.** Let \( H' \) be a graph, and let \( H \) be a \( \{2, 3, \ldots\} \)-subdivision of \( H' \). Then there exist \( t \in \mathbb{N} \) and \( \varepsilon > 0 \) such that for all \( c > 0 \) with \( c \leq 1/2 \), if \( G \) is \( \varepsilon \)-sparse and \( H \)-free, then \( G \) has a \( c \)-sparse \((\varepsilon c^T n, \varepsilon n)\)-pair.

**Proof.** By Lemma 3.5.18 it suffices to prove that \( H \) is almost bipartite. We first note that if \( H \) is a \( \{2, 3, \ldots\} \)-subdivision of a graph \( H \), then \( H \) is a \( \{2, 3\} \)-subdivision of some graph; therefore we may assume that \( H \) is a \( \{2, 3\} \)-subdivision of \( H' \).

Now let \( A = V(H') \subseteq V(H) \). It follows that \( A \) is a stable set in \( H \), and every vertex in \( H \setminus A \) has maximum degree one (since it is part of a path of length two
or three with ends in A). Since $H$ is a $\{2, 3\}$-subdivision of $H'$, it follows that $H$ is triangle-free, and therefore $H$ is almost bipartite. This implies the result of the lemma.

3.6 Fixed-size anticomplete pairs

In this section, we consider the following two questions.

**Question 3.6.1.** For which classes $\mathcal{C}$ is there a constant $c > 0$ such that for every tree $H \in \mathcal{C}$, there is an $\varepsilon > 0$ such that every $H$-free $\varepsilon$-sparse graph $G$ contains an anticomplete $(\varepsilon n^c, \varepsilon n^c)$-pair, where $n = |V(G)| > 1$?

**Question 3.6.2.** For which classes $\mathcal{C}$ is there a constant $c > 0$ such that for every tree $H \in \mathcal{C}$, there is an $\varepsilon > 0$ such that every $H$-free $\varepsilon$-sparse graph $G$ contains an anticomplete $(\varepsilon n^c, \varepsilon n)$-pair, where $n = |V(G)| > 1$?

Clearly, if Question 3.6.2 holds for a class of graphs, then so does Question 3.6.1.

Theorem 3.2.37 proves that if $\mathcal{C}$ is a set of forests, then Questions 3.6.1 and 3.6.2 hold with $c = 1$.

The following result shows that if $\mathcal{C}$ is the class of all graphs, then Question 3.6.1 is false.

**Lemma 3.6.3.** Let $c > 0$. Then there is a graph $H$ such that for every $\varepsilon > 0$, there exists an $H$-free $\varepsilon$-sparse graph with no anticomplete $(\varepsilon n^c, \varepsilon n^c)$-pair, where $n = |V(G)| > 1$.

**Proof.** Let $d = c/2$, let $k = 4 \lceil 1/d + 1 \rceil$, and let $H = K_k$. Let $\varepsilon > 0$, and let $p(n) = n^{-d}$. Let $G \sim G(n, p(n))$ be a random graph with $n$ vertices, and such that every edge is present independently with probability $p$. 

170
The probability of an anticomplete \((\varepsilon n^c, \varepsilon n^c)\)-pair is bounded by

\[
p_1 \leq (2n)^{2\varepsilon n^c} (1-p)^{(\varepsilon n^c)^2} \leq e^{2\varepsilon n^c \ln(2n) - n^{-d} \varepsilon^2 n^{2c} c} e^{2\varepsilon n^c \ln(2n) - \varepsilon^2 \frac{n^d}{2}} = e^{\varepsilon^c (2\varepsilon \ln(2n) - \varepsilon^2 n^{c/2})} \to 0,
\]
as \(n \to \infty\). The probability that \(G\) contains \(K_k\) is bounded by

\[
p_2 \leq (2n)^k p^{k(k-1)/2} = (2n \cdot n^{-d(k-1)/2})^k \leq (2n^{1-d(d/4)/2})^k = (n/2)^{-k} \to 0,
\]
as \(n \to \infty\). The expected degree of a vertex is \(2np = 2n^{1-d}\). It follows that for a given vertex, the probability that its degree is at least \(\varepsilon n/2\), is bounded by \(2n^{1-d}/(\varepsilon n/2) = 4n^{-d}/\varepsilon\), and thus the expected number of vertices of degree at least \(\varepsilon n/2\) is bounded by \(4n^{-d}/\varepsilon\). Let \(p_3\) be the probability that there are at least \(n\) vertices of degree at least \(\varepsilon n/2\); then \(p_3 \leq 4n^{1-d}/(\varepsilon n/2) \leq 8n^{-d}/\varepsilon \to 0\).

Therefore, for \(n\) sufficiently large, we have that \(\varepsilon n \geq 1\) and \(p_1 + p_2 + p_3 < 1\), and therefore, there is a graph \(G\) on \(2n\) vertices with no anticomplete \((\varepsilon n^c, \varepsilon n^c)\)-pair, no \(K_k\), and at most \(n\) vertices of degree at least \(\varepsilon n/2\); thus, if we delete all vertices of degree at least \(\varepsilon n/2\), we obtain an \(\varepsilon\)-sparse \(H\)-free graph with no anticomplete \((\varepsilon n^c, \varepsilon n^c)\)-pair. This concludes the proof.

\[\square\]

Lemma 3.6.3 does not imply that the class of trees (or forests) is the only class of graphs for which Questions 3.6.1 and 3.6.2 might hold, but the proof suggests that every such class has a bounded ratio between number of edges and number of vertices.

The following result is a weakening of Question 3.6.2 with a simple proof:

**Lemma 3.6.4.** For every tree \(T\), there is an \(\varepsilon > 0\) such that every \(\{T, K_3\}\)-free \(\varepsilon\)-sparse graph \(G\) contains an anticomplete \((\varepsilon n^{1/2}, \varepsilon n)\)-pair.
Proof. Let $T = T_{d,h}$ be a $d$-ary tree of height $h$. Since every tree is contained in a $d$-ary tree of height $h$ for some $d$ and $h$, it suffices to prove the lemma for such trees.

We will prove the following stronger statement:

\[\text{For all } d, h, \text{ there is an } \varepsilon(d, h) > 0 \text{ such that in every } \varepsilon(d, h)\text{-sparse }\]
\[\{T_{d,h}, K_3\}\text{-free graph } G \text{ with no anticomplete } (\varepsilon(d, h)n^{1/2}, \varepsilon(d, h)n)\text{-pair, for every } v \in V(G) \text{ with } d(v) \geq (\varepsilon(d, h)n^{1/2} + 1)d \text{, there is a } T_{d,h} \text{ with root } v \text{ in } G.\]

We prove this by induction on $h$. For $h = 1$, we let $\varepsilon(d, h) = 1$. It follows that every vertex $v \in V(G)$ with $d(v) \geq (\varepsilon(d, h)n^{1/2} + 1)d \geq d$ is the root of a $T_{d,h}$ in $G$.

Now let $h > 1$, and let $l = |V(T_{d,h-1})|$, let $\varepsilon = \varepsilon(d, h) = \varepsilon(d, h - 1)^2/(4 + 4d^2l + 2d)$. It follows that $\varepsilon \leq 1/(d + 2)$, and therefore, $\frac{1-2\varepsilon}{d} \geq \varepsilon$. Let $G, v$ be as in (3.20). We define $A_1, \ldots, A_d, B_0, B_1, \ldots, B_d, C_0, C_1, \ldots, C_d$ as follows with the following properties:

- for all $i \in \{1, \ldots, d\}$, $|N(A_i) \cap C_{i-1}| \geq ((1 - 2\varepsilon)/d - \varepsilon)n$;
- for all $i, j \in \{1, \ldots, d\}$ with $i < j$, $A_i$ is disjoint from $B_j \cup C_j \cup C_i$ and anticomplete to $C_j \cup C_i$;
- for all $i \in \{0, \ldots, d\}$, we have $|B_i| \geq (\varepsilon n^{1/2} + 1)(d - i)$ and $|C_i| \geq \left(1 - \frac{(1-2\varepsilon)i}{d} - \varepsilon\right)n$.

We let $B_0 = N(v)$, $C_0 = V(G) \setminus N[v]$. Now let $i \in \{1, \ldots, d\}$, and suppose that $B_{i-1}, C_{i-1}$ have been defined. Since $|B_{i-1}| \geq \varepsilon n^{1/2}$, it follows that $|C_{i-1} \setminus N(B_{i-1})| \leq \varepsilon n$, and therefore, $|N(B_{i-1}) \cap C_{i-1}| \geq |C_{i-1}| - \varepsilon n \geq \left(1 - 2\varepsilon - \frac{(1-2\varepsilon)i}{d}\right)n \geq \frac{1-2\varepsilon}{d}n$.

Now let $A_i \subseteq B_{i-1}$ be minimal with respect to inclusion subject to $|N(A_i) \cap C_{i-1}| \geq \left(\frac{1-2\varepsilon}{d} - \varepsilon\right)n$. By the minimality of $A_i$, it follows that $|N(A_i) \cap C_{i-1}| \leq \frac{1-2\varepsilon}{d}n$. We let $C_i = C_{i-1} \setminus N(A_i)$ and $B_i = B_{i-1} \setminus A_i$. It follows that $|C_i| \geq \left(1 - \frac{(1-2\varepsilon)i}{d} - \varepsilon\right)n \geq \varepsilon n$. Since $A_i$ is anticomplete to $C_i$ by construction, it follows that $|A_i| \leq \varepsilon n^{1/2}$, and so
\[ |B_i| \geq |B_{i-1}| - \varepsilon n^{1/2}, \] and therefore \[ |B_i| \geq (\varepsilon n^{1/2} + 1)(d - i). \] This proves that this construction yields sets with the desired properties.

Next, we construct trees \( T_1, \ldots, T_d \) such that for all \( i \in \{1, \ldots, d\} \), \( T_i \) is isomorphic to \( T_{d,h-1} \), the root of \( T_i \) is in \( A_i \), and the remaining vertices of \( T_i \) are in \( C_{i-1} \setminus C_i \), as follows. Let \( i \in \{1, \ldots, d\} \), and suppose that \( T_{i+1}, \ldots, T_d \) have been constructed. Let \( \mathcal{T} = V(T_{i+1}) \cup \cdots \cup V(T_d) \). Let \( A = A_i \) and \( C = C_{i-1} \setminus (C_i \cup N[\mathcal{T}]) \). By construction, \( A \) is disjoint from \( \mathcal{T} \) and anticomplete to \( \mathcal{T} \), since \( \mathcal{T} \subseteq C_i \cup N(v) \), and \( N(v) \) is stable.

We have that \( |C| \geq |N(A_i) \cap C_{i-1}| - d \varepsilon n \geq \left(\frac{1-2\varepsilon}{d} - \varepsilon(dl + 1)\right)n \geq \frac{1-\varepsilon}{d} n \), since \( \varepsilon \leq \frac{1}{4+2d^2 + 2d} \). Since \( A \) covers \( C \), it follows that there is a vertex \( v_i \in A \) with \( |N(a) \cap C| \geq \frac{1}{n} \frac{n}{2d^{1/2}} |n|^{1/2} \). Since \( G \) is \( \varepsilon \)-sparse, it follows that \( n \geq 1/\varepsilon \geq 1/\varepsilon(d, h - 1)^2 \).

It follows that \( \frac{1}{n} \frac{n}{2d^{1/2}} \geq 2d \varepsilon (d, h - 1)n^{1/2} \geq (\varepsilon(d, h - 1)n^{1/2} + 1)d \). We apply induction to \( C \cup \{v_i\} \). Since \( |C| \geq \frac{1}{2d} n \), it follows that with \( G' = G|\ (C \cup \{v_i\}) \), we have \( \varepsilon n \leq \varepsilon (d, h - 1)|V(G')| \), and so \( G' \) is \( \varepsilon (d, h - 1) \)-sparse, has no anticomplete \( (\varepsilon(d, h - 1)|V(G')|^{1/2}, \varepsilon(d, h - 1)|V(G')|\)-pair, and \( v_i \) has the required number of neighbors. It follows that \( T_i \) exists as claimed, with root \( v_i \). Moreover, \( V(T_i) \) is anticomplete to \( V(T_j) \) for all \( j > i \).

After constructing \( T_1, \ldots, T_d \), we let \( T = G|\ (\{v\} \cup V(T_1) \cup \cdots \cup V(T_d)) \); it follows that \( T \) is isomorphic to \( T_{d,h} \). This proves (3.20).

Now we let \( \varepsilon = \varepsilon(d, h)/(8d) \) with \( \varepsilon(d, h) \) as in (3.20). Since \( \{v\}, V(G) \setminus N[v] \) is an anticomplete pair for all \( v \in V(G) \), we may assume that \( \varepsilon n^{1/2} > 1 \). By Lemma 3.01 it follows that \( G \) has a vertex \( w \) of degree at least

\[
\frac{(1 - \varepsilon) n}{\varepsilon n^{1/2} + 1} \geq \frac{(1 - \varepsilon) n}{2 \varepsilon n^{1/2}} \geq \frac{1 - \varepsilon}{2 \varepsilon} \cdot n^{1/2} \geq \frac{1}{4 \varepsilon} n^{1/2} \geq 2d \varepsilon (d, h)n^{1/2} \geq (\varepsilon(d, h)n^{1/2} + 1) d.
\]

Now the result follows from (3.20) applied to \( G \) and \( w \).
3.7 Triangle-free graphs

The results of Section 3.5 and in particular of Section 3.5.3 resolve Question 3.3.2 for many graphs $H$, all of which are triangle-free. In this section, we present a result on anticomplete pairs in sparse triangle-free graphs. It is conjectured that Question 3.3.2 is true for triangles:

**Conjecture 3.7.1** (Conlon, Fox, Sudakov [18]). There exists $\varepsilon > 0$ such that every triangle-free graph $G$ contains an anticomplete $(\varepsilon n, \varepsilon n)$-pair, where $n = |V(G)|$.

Conjecture 3.7.1 is still open, and Lemma 3.7.2 is currently the best known bound $k$ for guaranteeing an anticomplete $(k, \varepsilon n)$-pair in a triangle-free graph.

**Lemma 3.7.2.** There exists an $\varepsilon > 0$ such that every triangle-free graph $G$ with $|V(G)| \geq 3$ contains an anticomplete $(\varepsilon \log n \log \log n, \varepsilon n)$-pair, where $n = |V(G)|$.

**Proof.** Let $G$ be a triangle-free graph. Then $N(v)$ is stable for every $v \in V(G)$, and thus we may assume that $\Delta(G) \leq \varepsilon \left(n + \frac{\log n}{\log \log n}\right)$. By choosing $\varepsilon \leq 1/8$, we may assume that $\Delta \leq n/4$ whenever $n \geq 8$. If $n < 8$, then choosing $\{u\}, \{v\}$ with $u, v$ non-adjacent yields the desired anticomplete pair.

Let $k \in \mathbb{N}, k \geq 3$ to be chosen later. Let $c = 1/4$, and suppose that $G$ does not contain an anticomplete $(k, cn)$-pair. Let $v_1, \ldots, v_k \in V(G)$. Then $|N(v_1) \cup N(v_2) \cup \cdots \cup N(v_k)| \geq n - cn - k$, since $\{v_1, \ldots, v_k\}, V(G) \setminus (N[v_1] \cup \cdots \cup N[v_k])$ is an anticomplete pair. Thus there exists a $v_i \in \{v_1, \ldots, v_k\}$ with $|N(v_i)| \geq (1 - c)n/k - 1$; let $w = v_i$. It follows that $|N(v_i)| \leq n/4 = cn$.

We claim that for all $t \in \mathbb{N}_0$ such that $n/(3k)^{t+1} \geq k$ and $t \leq k$, there exists a set $A_t$ with $|A_t| = t$ and $|C(A_t)| \geq n/(3k)^{t+1}$. Since $(1 - c)n/k - 1 \geq 3n/4k - 1 \geq n/3k$ (assuming $n/3k \geq k$), we may set $A_1 = \{w\}$.

Suppose that at step $t \leq k - 1$, we have constructed $A_t = \{w_1, \ldots, w_t\}$ with $cn \geq |C(A_t)| \geq n/(3k)^{t+1} \geq 3k^2$. Let $C = G \setminus (A_t \cup C(A_t))$. Suppose that there is a
set $B = \{v_1, \ldots, v_k\}$ of $k$ vertices in $C(A_t)$ satisfying $|N(v) \cap C| \leq (|C| - cn)/k$ for all $v \in B$. Then $|(N(v_1) \cup \cdots \cup N(v_k)) \cap C| \leq |C| - cn$, and consequently, $\{v_1, \ldots, v_k\}, C \setminus (N(v_1) \cup \cdots \cup N(v_k))$ contains an anticomplete $(k, cn)$-pair, a contradiction. It follows that all but at most $k-1$ vertices $v$ in $C(A_t)$ satisfy $|N(v) \cap C| \geq (|C| - cn)/k$, and thus $|E(C(A_t), C)| \geq (|C| - k) \cdot (|C| - cn)/|k|C|$. Consequently there is a vertex $w \in C$ with $|N(w) \cap C(A_t)| \geq (|C(A_t)| - k) \cdot (|C| - cn)/|k|C|$. Since $|C| = |V(G)| - t - |C(A_t)| \geq n - 2\Delta(G) \geq n/2 \geq 2cn$, it follows that $(|C| - cn)/|C| \geq 1/2$ and hence

$$\frac{(|C(A_t)| - k)(|C| - cn)}{k|C|} \geq |C(A_t)|/2k - 1/2 \geq |C(A_t)|/3k + k^2/6k - 1/2 \geq |C(A_t)|/3k.$$

We let $A_{t+1} = A_t \cup \{w\}$; it follows that $C(A_{t+1}) \geq N(w) \cap C(A_t) \geq |C(A_t)|/3k \geq n/(3k)^{t+2}$.

Now let $\varepsilon = 1/8$ and $k = \varepsilon \log_2 n / \log_2 \log_2 n$. Suppose that $\log_2 \log_2 n \geq 3\varepsilon$; then $3\varepsilon \log_2 n / \log_2 \log_2 n \leq \log_2 n$. It follows that

$$\log_2(n/(3k)^{k+2}) = \log_2 n - (k + 2) \log_2(3k)$$

$$= \log_2 n - (2 + \varepsilon \log_2 n / \log_2 \log_2 n) \log_2(3\varepsilon \log_2 n / \log_2 \log_2 n)$$

$$\geq \log_2 n - (2 + \varepsilon \log_2 n / \log_2 \log_2 n) \log_2 \log_2 n$$

$$\geq (1 - \varepsilon) \log_2 n - 2 \log_2 \log_2 n.$$

Suppose further that $n \geq 2^{16}$. Then $n \geq (\log_2 n)^4$, and thus $\log_2 n \geq 4 \log_2 \log_2 n$. It follows that $\log_2(n/(3k)^{k+2}) \geq 0$, and thus $n/(3k)^{k+1} \geq 3k$. Thus there exists a set $A_k$ of size $k$ with $|C(A_k)| \geq 3k$. Since $G$ is triangle-free, it follows that no vertex in $C = G \setminus (A_k \cup C(A_k))$ has a neighbor in $A_k$ and in $C(A_k)$, and thus either $(A_k, G \setminus N(A_k))$ or $(C(A_k), G \setminus N(C(A_k)))$ contains an anticomplete $(k, |C|/2)$-pair.
Since $|C| \geq n - 2\Delta(G)$, it follows that $|C|/2 \geq n/4 \geq \varepsilon n$. This is a contradiction, and thus the result follows in the case that $n \geq 2^{16}$ with $\varepsilon = 1/8$.

Now let $n \leq 2^{16}$. Then we let $\varepsilon = 1/16$; it follows that $\varepsilon \log_2 n/\log_2 \log_2 n \leq 1$. Since $\Delta(G) \leq n/4$, it follows that $G$ contains an anticomplete $(1, n/2)$-pair. Thus the lemma holds for all $G$ with $\varepsilon = 1/16$.

On the other hand, the following construction gives an upper bound on the best possible constant $\varepsilon$ for which every triangle-free graph $G$ contains a $(n^\varepsilon, cn)$ anticomplete pair for some constant $c$ that does not depend on $n$. The proof relies on the Lovász local lemma:

**Lemma 3.7.3** (Lovász Local Lemma [24]). Let $\mathcal{A}$ be a set of events and for $A \in \mathcal{A}$, let $Pr(A)$ denote the probability that $A$ occurs. If there is an assignment $x : \mathcal{A} \rightarrow [0, 1)$ such that for all $A \in \mathcal{A}$,

$$Pr(A) \leq x(A) \prod_{B \sim A} (1 - x(B)),$$

where the product is over all events $B \in \mathcal{A}$ such that $B$ is not independent from $A$, then the probability that none of the events in $\mathcal{A}$ occur is positive.

**Lemma 3.7.4.** For every $\varepsilon > 0$ there is a $C > 0$ such that for every sufficiently large $n$, there is an $n$-vertex triangle-free graph with no anticomplete $(Cn^{1/2}, \varepsilon n)$-pair.

**Proof.** We construct a graph $G$ on $n$ vertices by adding each edge at random with probability $p = 1/\sqrt{n}$. There are two kinds of events:

- for every triple $T$ of vertices in $G$, the event $A_T$ that $G|T$ is a triangle; and

- for every $(Cn^{1/2}, \varepsilon n)$-pair $X, Y$, the event $A_{X,Y}$ that the pair is anticomplete.

For every triple $T \subseteq V(G)$, the probability of $A_T$ is $p^3$; and for every $(Cn^{1/2}, \varepsilon n)$-pair $X, Y$, we have $Pr(A_{X,Y}) = (1 - p)^{C\varepsilon n^{3/2}} \leq e^{-pC\varepsilon n^{3/2}}$. 

176
For every triple $T \subseteq V(G)$, the number of other triples $T'$ such that $A_T$ and $A_{T'}$ are dependent is $n - 2$, since $T$ and $T'$ are dependent if and only if $|T \cap T'| \geq 2$. We bound the number of events $A_{X,Y}$ such that $A_T$ and $A_{X,Y}$ from above by $3^n$.

For every $(Cn^{1/2}, \varepsilon n)$-pair $X, Y$, the number of triples $T$ such that $A_{X,Y}$ and $A_T$ are dependent is $C \varepsilon n^{5/2}$, since $T \cap X, T \cap Y \neq \emptyset$ for such triples. Again, we bound the number of events $A_{X',Y'}$ such that $A_{X,Y}$ and $A_{X',Y'}$ are dependent by $3^n$.

We let $x(A_{X,Y}) = 3^{-n}$ and $x(A_T) = e^2 p^3$. Let $T$ be a triple. Then

$$x(A_T) \prod_{B \sim A_T} (1 - x(B)) \geq e^2 p^3 \cdot (1 - e^2 p^3)^{n-2} \cdot (1 - 3^{-n})^{3^n}$$

$$\geq e^2 p^3 \cdot e^{-e^2 p^3(n-2)} \cdot e^{-1} \geq p^3 = Pr(A_T)$$

if $n - 2 \leq e^{-2} p^{-3} = e^{-2} n^{3/2}$ which is true for $n$ sufficiently large.

Now let $X, Y$ be a $(Cn^{1/2}, \varepsilon n)$-pair. Then

$$x(A_{X,Y}) \prod_{B \sim A_{X,Y}} (1 - x(B)) \geq 3^{-n} (1 - e^2 p^3)^{C \varepsilon n^{5/2}} (1 - 3^{-n})^{3^n}$$

$$\geq 3^{-n} e^{-e^2 p^3 C \varepsilon n^{5/2}} e^{-1}.$$  

Thus our goal is to show that $-p \varepsilon Cn^{3/2} \leq -n - e^2 p^3 \varepsilon Cn^{5/2} - 1$. Clearly $-p/2 \cdot \varepsilon Cn^{3/2} \leq -n - 1$ for $C \geq 2/\varepsilon$; it remains to show that $-p/2 \cdot \varepsilon Cn^{3/2} \leq -e^2 p^3 \varepsilon n^{5/2}$. This is equivalent to showing that $-\varepsilon C \leq -2e^2 p^2 n = -2e^2$, which is true for $C \geq 2e^2/\varepsilon$. Therefore, Lemma 3.7.3 implies that the probability that none of these events occur is positive, i.e. there is a graph $G$ such that $G$ is triangle-free and has no anticomplete $(Cn^{1/2}, \varepsilon n)$-pair for all $n$ sufficiently large and for $C \geq 2e^2/\varepsilon$. This concludes the proof.

Since triangles satisfy the Erdős-Hajnal conjecture, solving Conjecture 3.7.1 in the affirmative would not imply an improved result for Conjecture 3.1.1. On the other
hand, if we intend to generalize the results of Section 3.5.3 then excluding a triangle is the simplest unsolved case that remains.
Chapter 4

Orthogonal strip systems in perfect graphs

Let $G$ be a graph. Recall that a graph $G$ is perfect if $\chi(H) = \omega(H)$ for every induced subgraph $H$ of $G$. We call a coloring $f$ of $G$ optimal if $|f(G)| = \chi(G)$.

The following definitions, in particular the definition of a $J$-strip system, were introduced in [11]. A graph $G$ is a double split graph if there exist $k, l \in \mathbb{N}$ such that $V(G) = \{a_1, \ldots, a_k, b_1, \ldots, b_k, c_1, \ldots, c_l, d_1, \ldots, d_l\}$, and

- for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $a_i$ is adjacent to $b_i$ and non-adjacent to $a_j$ and $b_j$;
- for all $i, j \in \{1, \ldots, l\}$ with $i \neq j$, $c_i$ is non-adjacent to $d_i$ and adjacent to $c_j$ and $d_j$;
- for all $i \in \{1, \ldots, k\}, j \in \{1, \ldots, l\}$, we have that $E(\{a_i, b_i\}, \{c_j, d_j\}) \in \{\{a_ic_j, b_id_j\}, \{a_id_j, b_ic_j\}\}$.

The line graph $L(G)$ is defined as the graph with $V(L(G)) = E(G)$; and $e, f \in E(G)$ are adjacent in $L(G)$ if $e \neq f$ and there exist $u, v, w \in V(G)$ such that $e = vu$, $f = vw$. A graph $G$ is basic if one of the following holds:
• $G$ or $G^c$ is bipartite;

• there exists a bipartite graph $H$ such that $G$ is isomorphic to $L(H)$ or $L(H)^c$; or

• $G$ is a double split graph.

Let $G$ be a graph. A 2-join in $G$ is a partition of $V(G)$ into $A_1, A_2, B_1, B_2, C_1, C_2$ such that

• $A_1$ is complete to $A_2$, and $B_1$ is complete to $B_2$;

• $A_1$ is anticomplete to $B_2 \cup C_2$, and $B_1$ is anticomplete to $A_2 \cup C_2$, and $C_1$ is anticomplete to $C_2$;

• for all $i \in \{1, 2\}$, every component of $G|(A_i \cup B_i \cup C_i)$ contains a vertex of $A_i$ and a vertex of $B_i$; and

• for all $i \in \{1, 2\}$, if $|A_i| = |B_i| = 1$ and $G|(A_i \cup B_i \cup C_i)$ is a path with one end in $A_i$ and the other end in $B_i$, then this path is odd and has length at least three.

A balanced skew partition of $G$ is a partition of $V(G)$ into non-empty sets $A, B, C, D$ such that $A$ is anticomplete to $D$, and $B$ is complete to $C$, and furthermore, for every two non-adjacent vertices in $B \cup C$, every induced path between them with interior in $A \cup D$ has even length, and for every two adjacent vertices in $A \cup D$, every induced antipath between them with interior in $B \cup C$ has even length.

The seminal strong perfect graph theorem \[11\] includes the following structural characterization of perfect graphs:

**Theorem 4.0.1** (Chudnovsky, Robertson, Seymour, Thomas \[11\]). Let $G$ be a perfect graph. Then either $G$ admits a balanced skew partition, or $G$ or $G^c$ admits a 2-join, or $G$ is basic.
The following result shows that perfect graphs can be colored in polynomial time:

**Theorem 4.0.2** (Grötschel, Lovász, Schrijver [28]). *There is a polynomial-time algorithm that finds an optimal coloring in a perfect graph.*

Theorem 4.0.2 predates Theorem 4.0.1 and relies on the ellipsoid method and semi-definite programming. Since the proof of Theorem 3.1.4 much effort has been devoted to finding a *combinatorial algorithm* for coloring perfect graphs, i.e. an algorithm that does not rely on the ellipsoid method. The following results are partial progress in this direction, and they suggest that balanced skew partitions are one of the main difficulties in this endeavor.

**Theorem 4.0.3** (Chudnovsky, Trotignon, Trunck, Vušković [17]). *There is a polynomial-time combinatorial algorithm for finding an optimal coloring in a perfect graph with no balanced skew partitions.*

**Theorem 4.0.4** (Chudnovsky, Lagoutte, Seymour, Spirkl [10]). *There is a combinatorial algorithm that returns an optimal coloring of a perfect graph \( G \) with running time \( O(|V(G)|^{\omega(G)+1})^2 \).*

The algorithm of Theorem 4.0.4 decomposes a given perfect graph recursively, using a balanced skew partition \( A, B, C, D \) and solving the coloring problem on \( G|_{(A \cup B \cup C)} \) and \( G|_{(B \cup C \cup D)} \). Since all vertices in \( B \cup C \) occur in both parts, bounding the size of the decomposition tree is one of the main difficulties in [10]. In Theorem 4.0.1 balanced skew partitions often arise from line graphs in the graph [11, Theorem 8.6]. Therefore, we are interested in balanced skew partitions that arise from certain line graphs, and in which \( G|_{(B \cup C)} \) contains certain line graphs. This leads to the study of orthogonal strip systems, defined below, which serve as a measure of “dimension” in a perfect graph. The main result of this section is a bound on the number of pairwise orthogonal strip systems in a perfect graph \( G \) in terms of the clique number \( \omega(G) \).
Let $k \in \mathbb{N}$. A graph $G$ is $k$-connected if $|V(G)| \geq k + 1$ and for every set $S$ of at most $k - 1$ vertices of $G$, $G \setminus S$ is connected. This implies that in a $k$-connected graph $G$, $|N(v)| \geq k$ for all $v \in V(G)$, for otherwise $G \setminus N(v)$ is not connected. A bipartite subdivision $H$ of a graph $G$ is a subdivision $H$ of $G$ such that $H$ is a bipartite graph.

For a graph $J$ and a graph $G$, an appearance $L(H)$ of $J$ in $G$ is an induced subgraph of $G$ isomorphic to $L(H)$, where $L(H)$ is the line graph of a bipartite subdivision of $J$. An appearance $L(H)$ of $J$ is non-degenerate if either $J = K_{3,3}$ and $H \neq K_{3,3}$; or if $J = K_4$ and $H$ does not contain four vertices of degree at least three in $H$ that form a 4-cycle; or if $J \not\in \{K_{3,3}, K_4\}$. An appearance $L(H)$ of $J$ in $G$ is overshadowed if there is a branch $B = B_e$ consisting of the path $u = x_0 - x_1 - \ldots - x_{l-1} - x_l = v$ for some $e = uv \in E(J)$ and for some odd $l \geq 3$, and there is a vertex $z \in V(G)$ such that $z$ is adjacent to every vertex of $G$ corresponding to an edge $uw$ or $vw$ in $E(H)$ with $w \not\in x_1, x_{l-1}$. For a graph $J$, a $J$-enlargement is graph $J'$ such that $J'$ is 3-connected and contains an induced subgraph isomorphic to a subdivision of $J$, and $J$ is not isomorphic to $J'$.

A graph is Berge if it contains no odd hole and no odd antihole as an induced subgraph. By Theorem 3.1.4, a graph is perfect if and only if it is Berge.

Let $G$ be perfect, and let $J$ be 3-connected. A $J$-strip system $(S, P)$ in a graph $G$ consists of

- for every edge $uv \in E(J)$, a set $S_{uv} = S_{vu} \subseteq V(G)$, called the strip corresponding to $uv$; and

- for every vertex $v \in V(J)$, a set $P_v \subseteq V(G)$, called the potato corresponding to $v$;

with the following definitions:

- for every edge $uv \in E(J)$, we let $S^*_{uv} = S_{uv} \setminus (P_u \cup P_v)$;

- for every vertex $v \in V(J)$, and for every $u \in N_J(v)$, we let $P_{vu} = P_v \cap S_{uv}$.
• \( V(S, P) = \bigcup_{uv \in E(J)} S_{uv} \);

satisfying the following conditions:

• \( \{ P_{vu} \}_{u \in N(v)} \) is a partition of \( P_v \) into non-empty subsets;

• the sets \( \{ S_{uv} \}_{uv \in E(J)} \) are pairwise disjoint;

• for every edge \( uv \in E(J) \), every vertex in \( S_{uv} \) is in a \( u-v \)-rung, i.e. an induced path with one end in \( P_{uv} \), one end in \( P_{vu} \), and interior in \( S_{uv}^{*} \);

• for every edge \( uv \in E(J) \), either every \( u-v \)-rung has odd length or every \( u-v \)-rung has even length;

• for every edge \( uv \in E(J) \), \( S_{uv}^{*} \) is anticomplete to \( V(S, P) \setminus (P_{uv} \cup P_{vu}) \);

• for \( u, v \in V(J) \) with \( u \neq v \), the set \( P_{u} \setminus P_{uv} \) is anticomplete to \( P_{v} \setminus P_{vu} \); and

• for \( v \in V(J) \) and \( u, w \in N_{J}(v) \) with \( u \neq w \), \( P_{vu} \) is complete to \( P_{vw} \).

A set \( \mathcal{R} = \{ R_{uv} \}_{uv \in E(J)} \) such that \( R_{uv} \) is a \( u-v \)-rung in \( S_{uv} \) for all \( uv \in E(J) \) is called a rung set. It follows from the above definitions that \( G \mid (\bigcup_{uv \in E(J)} V(R_{uv})) \) is an appearance \( L(H) \) of \( J \) for every rung set \( \{ R_{uv} \}_{uv \in E(J)} \). A \( J \)-strip system is non-degenerate if each such appearance is a non-degenerate appearance of \( J \). A \( J \)-strip system \( (S, P) \) is maximal if there is no \( J \)-strip system \( (S', P') \) with \( V(S, P) \subseteq V(S', P') \) and \( S_{uv}^{'} \cap V(S, P) = S_{uv} \) for all \( uv \in E(J) \), and \( P_{v}^{'} \cap V(S, P) = P_{v} \) for all \( v \in V(J) \).

For a \( J \)-strip system \( (S, P) \), we say that a set \( X \) seriously meets \( P \) if for all \( v \in V(J) \), there exists \( w \in N_{J}(v) \) such that for all \( u \in N_{J}(v) \setminus \{ w \} \), \( P_{vu} \subseteq X \). A vertex \( v \in V(G) \setminus V(S, P) \) is major for \( (S, P) \) if \( N_{G}(v) \) seriously meets \( P \). We say that \( X \subseteq V(S, P) \) is local with respect to \( (S, P) \) if either there is a vertex \( v \in V(J) \) such that \( X \subseteq P_{v} \), or there is an edge \( e \in E(J) \) such that \( X \subseteq S_{uv} \). A vertex \( v \in V(G) \setminus V(S, P) \) is minor for \( (S, P) \) if \( N_{G}(v) \) is local.
We say that a $J$-strip system $(S, P)$ in a graph $G$ is good if $J$ is 3-connected, $(S, P)$ is a non-degenerate and maximal $J$-strip system, and there is no $J$-enlargement with a non-degenerate appearance in $G$, and no overshadowed appearance of $J$ in $G$.

We need the following [11, Lemma 8.5]:

**Lemma 4.0.5** (Chudnovsky, Robertson, Seymour, Thomas [11]). Let $G$ be perfect, let $J$ be a 3-connected graph, and let $(S, P)$ a good $J$-strip system. Let $F \subseteq V(G) \setminus V(S, P)$ be connected, so that no member of $F$ is major with respect to $(S, P)$. Then the set of neighbors of $F$ in $V(S, P)$ is local.

We need another lemma, which is proved as part of (1) in [11, Lemma 8.6]:

**Lemma 4.0.6.** Let $G$ be perfect. Let $J$ be a 3-connected graph, and let $(S, P)$ be a good $J$-strip system. Let $M$ be the set of major vertices for $(S, P)$ and let $M'$ be a subset of $M$ with $G|M'$ anticonnected. Then the set of common neighbors of $M'$ seriously meets $P$.

### 4.1 Properties of orthogonal strip systems

Let $G$ be a graph, and let $J, J'$ be 3-connected. Let $(S^1, P^1)$ be a good $J$-strip system in $G$, and let $M^1$ be its set of major vertices. If $(S^2, P^2)$ is good $J'$-strip system in $G$, with major vertices $M^2$, we say that $(S^1, P^1)$ is orthogonal to $(S^2, P^2)$ if and only if $M^1$ seriously meets $P^2$ and $M^2$ seriously meets $P^1$.

Let $G$ be a graph. A vertex cover for $G$ is a set $X \subseteq V(G)$ such that for every edge $e = uv \in E(G)$, we have $\{u, v\} \cap X \neq \emptyset$. A matching in $G$ is a set $Y \subseteq E(G)$ such that for every $v \in V(G)$, there exists at most one $w \in V(G)$ such that $vw \in Y$.

We use König’s theorem:

**Theorem 4.1.1** (König [32]). Let $G$ be a bipartite graph; let $X$ be a vertex cover for $G$ with $|X|$ minimum, and let $Y$ be a matching in $G$ with $|Y|$ maximum. Then $|X| = |Y|$. 

184
The following lemma is used to establish a simple criterion for orthogonality. Let $G$ be a perfect graph, let $J$ be 3-connected, and let $(S, P)$ be a good $J$-strip system. We call $x_1, x_2, x_3$ a spanning triple for $(S, P)$ if for all $v \in V(J)$, we have $|P_v \cap \{x, y, z\}| \leq 1$; and if $\{x_1, x_2, x_3\} \subseteq \bigcup_{v \in V(J)} P_v$.

**Lemma 4.1.2.** Let $G$ be a perfect graph, let $J$ be 3-connected, and let $(S, P)$ be a good $J$-strip system in $G$. Let $S' \subseteq \bigcup_{v \in V(J)} P_v$. Then the following are equivalent:

- $S'$ contains three vertices $x_1, x_2, x_3$ that are a spanning triple for $(S, P)$;
- for all $v, u \in V(J)$, $S' \setminus (P_u \cup P_v) \neq \emptyset$.

**Proof.** If there exist $v, u \in V(J)$ such that $S' \subseteq P_u \cup P_v$, then certainly $x, y, z$ as in the first bullet do not exist. Now suppose that the second bullet holds; our goal is to prove that the first bullet holds as well.

Let $\{R_{uv}\}_{uv \in E(J)}$ be a rung set, and let $\{S'_{uv}\}_{uv \in E(J)}$ be a collection of sets, where for all $uv \in E(J)$, $R_{uv}$, $S'_{uv}$ are chosen as follows:

- if $S' \cap S_{uv} = \emptyset$, we let $R_{uv}$ be an arbitrary $u$-$v$-rung; and we let $S'_{uv} = \emptyset$;
- if there is a vertex $s \in S' \cap ((P_{uv} \setminus P_{vu}) \cup (P_{vu} \setminus P_{uv}))$, we let $R_{uv}$ be a $u$-$v$-rung containing $s$ (and therefore $R_{uv}$ has length at least one); we let $S'_{uv} = P_{uv} \cap V(R_{uv})$ if $P_{vu} \setminus P_{uv} = \emptyset$, and $S'_{uv} = P_{vu} \cap V(R_{uv})$ if $P_{uv} \setminus P_{vu} = \emptyset$, and $S'_{uv} = (P_{uv} \cup P_{vu}) \cap V(R_{uv})$ otherwise.
- otherwise, it follows that there is a vertex $s \in (S' \cap S_{uv}) \subseteq P_{uv} \cap P_{vu}$; and we let $R_{uv} = s$, $S'_{uv} = \{s\}$.

The second bullet ensures that if $R_{uv}$ is chosen according to the second bullet, then $S'_{uv} \cap P_{uv} \neq \emptyset$ if and only if $S' \cap (P_{uv} \setminus P_{vu}) \neq \emptyset$. Furthermore, this construction ensures that if $S'_{uv} \cap S_{uv} \neq \emptyset$, then $R_{uv}$ consists of a single vertex if and only if $(S' \cap S_{uv}) \subseteq P_{uv} \cap P_{vu}$.

185
Let $H$ be a bipartite subdivision of $J$ such that $L(H)$ is isomorphic to $G|\bigcup_{uv \in E(J)} V(R_{uv})$. We let $S^* = \bigcup_{uv \in E(J)} S'_{uv}$. Let $S^{**}$ be the set of vertices in $L(H)$ corresponding to vertices of $S^*$. Let $E^*$ be the set of edges in $E(H)$ corresponding to vertices in $S^{**} \subseteq V(L(H))$.

It follows that $H$ is bipartite. Let $H^*$ arise from $H$ by replacing every branch of length two by a branch of length four. Then $H^*$ is a bipartite subdivision of $J$, and we map $E(H)$ to a subset of $E(H^*)$ by mapping the edges $ab, bc$ of a branch $B_{ac} = a-b-c$ in $H$ to the edges $a-x$ and $z-c$ of the new branch $B^*_{a,c} = a-x-y-z-c$ in $H^*$, and by the identity for all other edges. Let $H' = (V(H^*), E^*)$. Since $H'$ is a subgraph of $H^*$, it follows that $H'$ is bipartite as well. By König’s theorem (Theorem 4.1.1), either $H'$ has a matching of size at least three or a vertex cover of size at most two. Suppose first that $H'$ has a matching $Y = \{y_1, y_2, y_3\}$ of size three. The set $Y$ corresponds to a set $Y'$ of vertices in $S^{**}$, which in turn corresponds to a set $Y^*$ of vertices in $S^*$. Let $(u_1, v_1), (u_2, v_2), (u_3, v_3)$ with $u_i, v_i \in V(J)$ such that $y_i \in Y^* \cap P_{u_iv_i}$ for all $i \in \{1, 2, 3\}$. By construction, for all $i \in \{1, 2, 3\}$ there exists an $z_i \in V(H')$ such that $y_i$ is an edge of the branch $B^*_{u_iv_i}$ of the subdivision $H^*$ of $J$, and $y_i = u_i z_i$. Since $Y$ is a matching, it follows that $u_1, u_2, u_3$ are distinct.

By the definition of $S^*$, it follows that $S' \cap P_{u_iv_i} \neq \emptyset$ for all $i \in \{1, 2, 3\}$; we let $x_i \in S' \cap P_{u_iv_i}$ for all $i \in \{1, 2, 3\}$, where we let $x_i \in S' \cap (P_{u_iv_i} \setminus P_{v_iu_i})$ if possible.

Suppose that there exists a $v \in V(J)$ such that $|\{x_1, x_2, x_3\} \cap P_v| > 1$. It follows that $v \in \{u_1, u_2, u_3, v_1, v_2, v_3\}$. By symmetry, we may assume that $x_1, x_2 \in P_v$. It follows that $v \in \{u_i, v_i\}$ for all $i \in \{1, 2\}$. Since $u_1 \neq u_2$, by symmetry, we may assume that $v = v_2$. Since $x_2 \in P_{v_2}$, it follows that $S' \cap (P_{u_2v_2} \setminus P_{v_2u_2}) = \emptyset$. Since $u_2 z_i \in E^*$, and $u_i z_i$ corresponds to a vertex in $P_{uv}$, but $S' \cap (P_{u_2v_2} \setminus P_{v_2u_2}) = \emptyset$, it follows that $R_{u_2v_2}$ was chosen according to the third bullet above. Therefore, $S' \cap S_{u_2v_2} \subseteq P_{u_2} \cap P_{v_2}$, and so $R_{u_2v_2}$ consists of a single vertex, and so the branch $B^*_{u_2v_2}$ has length one. This implies that $y_2 = u_2 v_2$.  

186
Suppose first that \( v = v_1 \). Then the same argument as above shows that \( y_1 = u_1v_1 \), and so \( v = v_1 = v_2 \) is incident with two different edges of \( Y \), contrary to the definition of a matching. It follows that \( v = u_1 \). But then \( y_1 = u_1z_1 \), and again \( v = v_2 = u_1 \) is incident with two different edges of \( Y \), a contradiction. This shows that \( | \{ x_1, x_2, x_3 \} \cap P_v | \leq 1 \) for all \( v \in V(J) \), and so the statement of the lemma holds.

It remains to consider the second case, that \( H' \) has a vertex cover \( X \) with \( |X| \leq 2 \). Let \( X = \{ v_1, v_2 \} \), where possibly \( v_1 = v_2 \). We first show that we may assume \( v_1, v_2 \in V(J) \). Suppose that \( v_1 \notin V(J) \). It follows that \( v_1 \) is an interior vertex of a branch \( B_e^* \) of \( H^* \) for some edge \( e = uw \in E(J) \). Let \( P_e = x_1 \ldots x_l \). By construction, it follows that \( E^* \cap E(P_e) \subseteq \{ x_1x_2, x_l-1, x_l \} \). Since \( P_e \) has an interior vertex, and \( P_e \) does not have length two, it follows that \( x_2 \neq x_{l-1} \). Therefore, \( v_1 \) is adjacent to at most one edge of \( E^* \). But then \( \{ x_1, v_2 \} \) or \( \{ x_1, v_2 \} \) is a vertex cover of \( H' \) as well.

Now suppose that \( v_1, v_2 \in V(J) \). We claim that \( S' \subseteq P_{v_1} \cup P_{v_2} \). Suppose not, and let \( s \in S' \setminus (P_{v_1} \cup P_{v_2}) \). Let \( u, v \in V(J) \) such that \( s' \in P_{uv} \). It follows that \( u \notin \{ v_1, v_2 \} \). Then there is an edge \( ux \in E(H') \) for some \( x \in V(H') \). Since \( \{ v_1, v_2 \} \) is a vertex cover of \( H' \), it follows that \( x \in v_1v_2 \). By symmetry, we may assume that \( x = v_1 \). Since the branch \( B_{uv_1}^* \) consists of a single edge, it follows that \( R_{uv_1} \) consists of a single vertex. But this implies that \( S' \cap S_{uv_1} \subseteq (P_{uv_1} \cap P_{v_1u}) \), and therefore \( s' \in P_{v_1u} \subseteq P_{v_1} \), a contradiction. This implies that \( S' \subseteq P_{v_1} \cup P_{v_2} \). This contradicts our assumption, and hence the lemma is proved.

**Lemma 4.1.3.** Let \( G \) be a perfect graph, let \( J \) be 3-connected, and let \( (S, P) \) be a good \( J \)-strip system in \( G \). Suppose that \( M \subseteq V(S, P) \) seriously meets \( P \). Then \( M \) contains vertices \( x_1, x_2, x_3 \) that form a spanning triple in \( (S, P) \). Moreover, all common neighbors of \( x_1, x_2, x_3 \) are major for \( (S, P) \), and \( G \mid (M \cap V(S, P)) \) is anticonnected.

**Proof.** Suppose first that \( M \) does not contain a spanning triple in \( (S, P) \). By Lemma 4.1.2 it follows that there are two vertices \( v, u \in V(J) \) such that \( M \subseteq P_v \cup P_u \). Since \( M \) seriously meets \( P \), it follows that for every \( w \in V(J) \setminus \{ v, u \} \), we have \( M \cap P_w \subseteq \)

187
$P_u \cup P_v$, and therefore $|N_J(w) \setminus \{u, v\}| \leq 1$. Consequently, every connected component of $J \setminus \{u, v\}$ has size at most two. Since $J$ is 3-connected, it follows that $J \setminus \{u, v\}$ is connected, and therefore $|V(J)| = 4$. Let $V(J) = \{u, v, w, z\}$. It follows that $J = K_4$, and $P_{wu}, P_{zu} \subseteq P_u$, and $P_{uv}, P_{zu} \subseteq P_v$. Let $\mathcal{R} = \{R_{xy}\}_{xy \in E(J)}$ be a rung system for $(S, P)$. It follows that $R_{wu}, R_{zu}, R_{uv}, R_{zu}$ consist of single vertices. Let $L(H)$ be the appearance of $J$ corresponding to $\mathcal{R}$. Then the branches $B_{wu}, B_{zu}, B_{uv}, B_{zu}$ of $H$ all have length one, and hence they form a four-cycle. This implies that $L(H)$ is a degenerate appearance of $J$ in $G$, contrary to the assumption that $(S, P)$ is a good $J$-strip system. It follows that $M$ contains a spanning triple $x_1, x_2, x_3$ for $(S, P)$.

Now let $x_1, x_2, x_3 \in M \cap V(S, P)$ be a spanning triple, and let $v_1, v_2, v_3 \in V(J)$ be distinct such that $x_i \in P_{v_i}$ for all $i \in \{1, 2, 3\}$. Suppose that $x_1 x_2 \in E(G)$. Then $x_1, x_2 \in P_{v_1 v_2} \cup P_{v_2 v_1}$. By the properties of a $J$-strip system, it follows that $x_1, x_2$ are anticomplete to $P_{v_3} \setminus (P_{v_3 v_1} \cup P_{v_3 v_2})$. Since $x_3 \notin P_{v_1} \cup P_{v_2}$, this implies that $|E(G) \{x_1, x_2, x_3\}| \leq 1$, and therefore $G \{x_1, x_2, x_3\}$ is anticonnected.

Now let $x$ be a common neighbor of $x_1, x_2, x_3$, and suppose that $N(x)$ is local. Since there is no $v \in V(J)$ with $\{x_1, x_2, x_3\} \subseteq P_v$ by the definition of a spanning triple, it follows that there exists an edge $uv \in E(J)$ with $\{x_1, x_2, x_3\} \subseteq S_{uv}$. But then either $|P_v \cap \{x_1, x_2, x_3\}| > 1$ or $|P_u \cap \{x_1, x_2, x_3\}| > 1$, a contradiction. This implies that every common neighbor of $x_1, x_2, x_3$ is major.

This implies that $G(M \cap V(S, P))$ is anticonnected: $G \{x_1, x_2, x_3\}$ is anticonnected, and every other vertex in $M \cap V(S, P)$ has a non-neighbor in $\{x_1, x_2, x_3\}$. □

Lemma 4.1.4. Let $G$ be a perfect graph, let $J, J'$ be 3-connected, let $(S^1, P^1)$ be a good $J$-strip system in $G$, and let $(S^2, P^2)$ be a $J'$-good strip system in $G$. Then $(S^1, P^1)$ is orthogonal to $(S^2, P^2)$ if and only if there is a spanning triple of $(S^1, P^1)$ consisting of vertices $x_1, x_2, x_3$ that are major for $(S^2, P^2)$.

Proof. Suppose that $(S^1, P^1)$ is orthogonal to $(S^2, P^2)$. Let $M^2$ be the set of major vertices for $(S^2, P^2)$. By the definition of orthogonality, it follows that $M^2$ seriously
meets $P^1$, and therefore, by Lemma 4.1.3, $M^2$ contains a spanning triple of $(S^1, P^1)$. This proves one direction.

For the other direction, we let $x_1, x_2, x_3$ be a spanning triple for $(S^1, P^1)$ and suppose that $x_1, x_2, x_3$ are major for $(S^2, P^2)$. Let $M^1$ be the set of major vertices for $(S^1, P^1)$ and let $M^2$ be the set of major vertices for $(S^2, P^2)$. By Lemma 4.1.3, $G\{|x_1, x_2, x_3\}$ is anticonnected. By Lemma 4.0.6, it follows that the set $C$ of common neighbors of $\{x_1, x_2, x_3\}$ seriously meets $P^2$. By Lemma 4.1.3, it follows that $C \subseteq M^1$, and therefore $M^1$ seriously meets $P^2$. Since $C$ seriously meets $P^2$, again by Lemma 4.1.3, it follows that $C$ contains a spanning triple $\{y_1, y_2, y_3\}$ for $(S^2, P^2)$. By Lemma 4.1.3, $G\{|y_1, y_2, y_3\}$ is anticonnected. By Lemma 4.0.6, it follows that the set $C'$ of common neighbors of $\{y_1, y_2, y_3\}$ seriously meets $P^1$. By Lemma 4.1.3, it follows that $C \subseteq M^2$, and therefore $M^2$ seriously meets $P^1$. This proves that $(S^1, P^1)$ and $(S^2, P^2)$ are orthogonal.

Let $G$ be a perfect graph, let $J$ be 3-connected, let $(S, P)$ be a good $J$-strip system in $G$. Let $X \subseteq V(G)$. We let $N^P(X)$ denote the set of vertices $x$ of $V(G)$ such that $x$ is complete to $X$, and $x \in P_v$ for some $v \in V(J)$.

**Lemma 4.1.5.** Let $G$ be a perfect graph, let $J, J'$ be 3-connected, let $(S^1, P^1)$ be a good $J$-strip system in $G$, and let $(S^2, P^2)$ be a $J'$-good strip system in $G$. Suppose that $(S^1, P^1)$ is orthogonal to $(S^2, P^2)$. For $i \in \{1, 2\}$, we let $M^i$ be the set of major vertices for $(S^i, P^i)$.

Then, for $i \in \{1, 2\}$, there exists an anticomponent $M_i \subseteq G\{|M^i\}$ such that the sets $X_{12} = N^{P^1}(V(M_1))$, $X_{21} = N^{P^2}(V(M_2))$ are complete to each other. Additionally, $X_{12} \subseteq M_2$ and $X_{21} \subseteq M_1$.

Moreover, $M^2 \setminus V(M_2) \subseteq M^1$ and $M^1 \setminus V(M_1) \subseteq M^2$.

**Proof.** By Lemma 4.1.3, it follows that $G\{|Y_1\}$ is anticonnected, where $Y_1 = M^1 \cap V(S^2, P^2)$, and therefore there is an anticomponent $M_1$ of $G\{|M^1\}$ such that $Y_1 \subseteq
\( V(M_1) \). Let \( X_{12} = N^{P^1}(V(M_1)) \). By Lemma 4.0.6 it follows that \( X_{12} \) seriously meets \( P^1 \).

Since \( X_{12} \) is complete to \( Y_1 \), and since \( Y_1 \) seriously meets \( P^2 \), it follows from Lemma 4.1.3 that \( G|X_{12} \) is anticonnected and \( X_{12} \subseteq M^2 \). It follows that there is an anticomponent \( M_2 \) of \( G|M^2 \) such that \( X_{12} \subseteq V(M_2) \). Let \( X_{21} = N^{P^2}(V(M_2)) \). It follows that \( X_{21} \) is complete to \( X_{12} \subseteq V(M_2) \).

The final part of the lemma follows because for all \( i \in \{1, 2\} \), the set \( M_i \setminus V(M_i) \) is complete to \( V(M_i) \), and complete to \( X_{(3-i)i} \). Since \( X_{(3-i)i} \) seriously meets \( P^{3-i} \) by Lemma 4.0.6 it follows from Lemma 4.1.3 that \( M_i \setminus V(M_i) \subseteq M^{3-i} \) for all \( i \in \{1, 2\} \).

4.2 Finding a clique

The goal of this section is to show that if \( G \) is perfect and contains \( k \) pairwise orthogonal good strip systems, then \( k \leq 2\omega(G) \).

**Lemma 4.2.1.** Let \( G \) be a perfect graph, and let \( J_1, \ldots, J_k \) be 3-connected graphs. For \( i \in \{1, \ldots, k\} \), we let \((S^i, P^i)\) be a good \( J_i \)-strip system in \( G \) with major vertices \( M^i \). Suppose that \((S^1, P^1), \ldots, (S^k, P^k)\) are pairwise orthogonal.

For all \( i, j \in \{1, \ldots, k\} \) with \( i \neq j \), we define \( X_{ij} = N^{P^i}(V(M_i)) \), where \( M_i \) is an anticomponent of \( M^i \) containing \( M^i \cap V(S^j, P^j) \).

Then the sets \( Z_i = \cap_{j \neq i} X_{ij} \) are pairwise disjoint and complete to each other.

**Proof.** By Lemma 4.1.5 it follows that for all \( i, j \in \{1, \ldots, k\} \) with \( i \neq j \), the sets \( X_{ij} \) and \( X_{ji} \) are disjoint and complete to each other. Since \( Z_i \subseteq X_{ij} \) and \( Z_j \subseteq X_{ji} \), it follows that \( Z_i \) and \( Z_j \) are disjoint and complete to each other.

We will prove that using the notation of Lemma 4.2.1 at least half of the \( Z_i \) are non-empty.
We begin by finding a way of ordering strip systems. Let $G$ be perfect, let $J$ be 3-connected, and let $(S, P)$ be a good $J$-strip system in $G$. A set $Y \subseteq V(G)$ is a serious $P$-set if $Y \subseteq \bigcup_{v \in V(J)} P_v$ and $Y$ seriously meets $P$.

Let $M$ be the set of major vertices for $(S, P)$ in $G$. We define a function $f_{(S, P)}$ as follows. Let $Y$ be a serious $P$-set. Then $f_{(S, P)}(Y)$ is defined to be the number of vertices $y \in M$ such that $y$ is in an anticomponent $M'$ of $G|M$, and $V(M')$ is complete to $Y$.

Let $J_1, \ldots, J_k$ be 3-connected graphs. For $i \in \{1, \ldots, k\}$, we let $(S^i, P^i)$ be a good $J_i$-strip system in $G$ with major vertices $M^i$. Suppose that $(S^1, P^1), \ldots, (S^k, P^k)$ are pairwise orthogonal. Then $(S^1, P^1), \ldots, (S^k, P^k)$ is an increasing ordering if for all $i \in \{1, \ldots, k\}$, there exists a serious $P^i$-set $Y_i$ such that for all $j \in \{i + 1, \ldots, k\}$, and for every serious $P^j$-set $Y_j$, $f_{(S^i, P^i)}(Y_i) \geq f_{(S^j, P^j)}(Y_j)$.

**Lemma 4.2.2.** Let $J_1, \ldots, J_k$ be 3-connected graphs. For $i \in \{1, \ldots, k\}$, we let $(S^i, P^i)$ be a good $J_i$-strip system in $G$ with major vertices $M^i$. Suppose that $(S^1, P^1), \ldots, (S^k, P^k)$ are pairwise orthogonal. Then there is an increasing ordering of $(S^1, P^1), \ldots, (S^k, P^k)$.

**Proof.** We prove this by induction on $k$. For $k = 1$, $(S^1, P^1)$ is an increasing ordering. Now let $k > 1$, and let $i \in \{1, \ldots, k\}$ such that $i$ maximizes $\max \{f_{(S^i, P^i)}(Y) : Y \text{ is a serious } P^i\text{-set}\}$. By symmetry, we may assume that $i = 1$. By induction, $(S^2, P^2), \ldots, (S^k, P^k)$ has an increasing ordering; by symmetry, we may assume that this ordering is $(S^2, P^2), \ldots, (S^k, P^k)$. Now $(S^1, P^1), \ldots, (S^k, P^k)$ is an increasing ordering.

**Lemma 4.2.3.** Let $G$ be a perfect graph, and let $J_1, \ldots, J_k$ be 3-connected graphs. For $i \in \{1, \ldots, k\}$, we let $(S^i, P^i)$ be a good $J_i$-strip system in $G$. Suppose that $(S^1, P^1), \ldots, (S^k, P^k)$ is an increasing ordering of pairwise orthogonal strip systems. For $i \in \{1, \ldots, k\}$, we let $M^i$ denote the set of major vertices for $(S^i, P^i)$ in $G$. 191
For all \(i, j \in \{1, \ldots, k\}\) with \(i \neq j\), we define \(X_{ij} = N_{P^i}(V(M_{ij}))\), where \(M_{ij}\) is an anticomponent of \(M^i\) containing \(M^i \cap V(S^j, P^j)\). For \(i \in \{1, \ldots, k\}\), we let \(Y_i\) be a serious \(P^i\)-set with \(f_{(S^i, P^i)}(Y_i)\) maximum.

For all \(i, j \in \{1, \ldots, k\}\), if \(j > i\), then \(X_{ij}\) contains \(Y_i\).

**Proof.** Suppose not, and let \(i < j\) such that \(X_{ij}\) does not contain \(Y_i\). By Lemma 4.1.5 it follows that \(X_{ji} \subseteq V(M_{ij})\) Let \(M^*\) be the union of the vertex sets of the anticomponents of \(G|M^i\) that are complete to \(Y_i\). Since \(X_{ij}\) does not contain \(Y_i\), it follows that \(V(M_{ij})\) is not complete to \(Y_i\), and therefore \(f_{(S^i, P^i)}(Y_i) = |M^*| \leq |M^i \setminus V(M_{ij})|\). By definition, \(V(M_{ji})\) is complete to \(X_{ji}\). By Lemma 4.1.5 it follows that \(M^* \subseteq M^i\), and that \(M^j \setminus M^i \subseteq V(M_{ji})\), and \(M^j \setminus V(M_{ji}) \subseteq M^i\).

We claim that \(M^* \cup V(M_{ji})\) is complete to \(M^j \setminus (M^* \cup V(M_{ji}))\). Suppose not; let \(m \in M^j \setminus (M^* \cup V(M_{ji}))\) such that \(m\) has a non-neighbor in \(M^* \cup V(M_{ji})\). Since \(M_{ji}\) is an anticomponent of \(G|M^j\), it follows that \(m\) has a non-neighbor in \(M^*\). It follows that \(m \in M^j \setminus V(M_{ji}) \subseteq M^i\). Since \(M^*\) is the union of vertex sets of anticomponents of \(G|M^i\), it follows that \(m\) is complete to \(M^*\). This is a contradiction, and our claim is proved.

It follows \(M^* \cup V(M_{ji})\) is the vertex set of a union of anticomponents of \(G|M^j\). Moreover, \(V(M_{ji})\) is complete to \(X_{ji}\), and \(M^*\) is complete to \(V(M_{ij})\). Since \(X_{ji} \subseteq V(M_{ij})\), it follows that \(M^* \cup V(M_{ji})\) is complete to \(X_{ji}\).

Since \(X_{ji}\) is a serious \(P^j\)-set by Lemma 4.1.5 and Lemma 4.0.6, it follows that

\[
f_{(S^j, P^j)}(X_{ji}) \geq |M^*| + |V(M_{ji})| > |M^*| = f_{(S^i, P^i)}(Y_i).
\]

Lemma 4.2.3 shows that \(Y_i \subseteq \bigcap_{j > i} X_{ij}\), which proves that \(\bigcap_{j > i} X_{ij}\) is a serious \(P^i\)-set. It remains to consider \(X_{ij}\) with \(i > j\). We start with the following lemma.

**Lemma 4.2.4.** Let \(G\) be a perfect graph, and let \(J_1, \ldots, J_k\) be 3-connected graphs. For \(i \in \{1, \ldots, k\}\), we let \((S^i, P^i)\) be a good \(J_i\)-strip system in \(G\). Suppose that
(S^1, P^1), \ldots, (S^k, P^k) is an increasing ordering of pairwise orthogonal strip systems. For i ∈ \{1, \ldots, k\}, we let M_i denote the set of major vertices for (S^i, P^i) in G.

For all i, j ∈ \{1, \ldots, k\} with i ≠ j, we define X_{ij} = N^{P^i}(V(M_{ij})), where M_{ij} is an anticomponent of M^i containing M^i ∩ V(S^j, P^j). For i ∈ \{1, \ldots, k\}, we let Y_i be a serious P^i-set with f_{(S^i, P^i)}(Y_i) maximum.

For all i, j, p ∈ \{1, \ldots, k\}, with i, j, p distinct, if X_{ip} ⊈ X_{ij}, then X_{pj} = X_{pi}.

Proof. Suppose that X_{ip} ≠ X_{ij}. It follows that M_{ij} ≠ M_{ip}. Since X_{ji} ⊆ V(M_{ij}) and X_{pi} ⊆ V(M_{ip}) by Lemma \textbf{4.1.5}, it follows that X_{pi} is complete to X_{ji}. Since M_{ij}, M_{ip} are different anticomponents of M^i, it follows that V(M_{ij}) ⊆ M^i \setminus V(M_{ip}). Furthermore, by Lemma \textbf{4.1.5} it follows that M^i \setminus V(M_{ip}) ⊆ M^p, which implies that since X_{ji} ⊆ V(M_{ij}), we have X_{ji} ⊆ M^p. By the definition of M_{pi}, it follows that X_{ip} ⊆ V(M_{pi}) ⊆ M^p. By Lemma \textbf{4.0.6} it follows that X_{ji} is a serious P^j-set and X_{ip} is a serious P^i-set, and hence, by Lemma \textbf{4.1.3} G|X_{ji} and G|X_{ip} are anticonnected.

We consider two cases. Suppose first that X_{ji} ⊆ V(M_{pi}). Since every subset of V(S^j, P^j) containing a serious P^j-set is anticonnected by Lemma \textbf{4.1.3} and since X_{ji} ⊆ V(S^j, P^j), it follows that V(S^j, P^j) ∩ M^p ⊆ V(M_{pi}). It follows that X_{jp} ⊆ V(S^j, P^j) ∩ M^p ⊆ V(M_{pi}), and therefore M_{pi} = M_{pj}. But then X_{pi} = N^{P^p}(V(M_{pi})) = N^{P^p}(V(M_{pj})) = X_{pj}, as claimed.

Therefore, we may assume that X_{ji} ⊈ V(M_{pi}). Let M'_{ij} be the anticomponent of M^p containing V(M_{ij}) ⊆ M^p. It follows that X_{ji} ⊆ V(M'_{ij}). It follows that V(M'_{ij}) is complete to V(M_{pi}). In particular, it follows that V(M_{ij}) is complete to X_{ip} ⊆ V(M_{pi}). Since X_{ij} = N^{P^i}(V(M_{ij})), it follows that X_{ip} ⊆ X_{ij}. This is a contradiction, and the statement of the lemma follows.

\textbf{Lemma 4.2.5.} Let G be a perfect graph, and let J_1, \ldots, J_k be 3-connected graphs. For i ∈ \{1, \ldots, k\}, we let (S^i, P^i) be a good J_i-strip system in G. Suppose that (S^1, P^1), \ldots, (S^k, P^k) is an increasing ordering of pairwise orthogonal strip systems. For i ∈ \{1, \ldots, k\}, we let M^i denote the set of major vertices for (S^i, P^i) in G.
For all $i, j \in \{1, \ldots, k \}$ with $i \neq j$, we define $X_{ij} = N^{pi}(V(M_{ij}))$, where $M_{ij}$ is an anticomponent of $M^i$ containing $M^i \cap V(S^j, P^j)$. For $i \in \{1, \ldots, k \}$, we let $Y_i$ be a serious $P^i$-set with $f_{(S^i, P^i)}(Y_i)$ maximum.

For each $j \in \{1, \ldots, k \}$, there is at most one $i \in \{j + 1, \ldots, k \}$ such that for all $p \in \{j + 1, \ldots, k \} \setminus \{i \}$, $X_{ip} \not\subseteq X_{ij}$.

Consequently, there are at most $k$ pairs $(i, j)$ with $i, j \in \{1, \ldots, k \}$ such that $i < j$ and for all $p \in \{j + 1, \ldots, k \} \setminus \{i \}$, $X_{ip} \not\subseteq X_{ij}$.

Proof. Suppose not. Let $j, i, h \in \{1, \ldots, k \}$ with $j < i < h$ such that for all $p \in \{j + 1, \ldots, k \}$, if $p \neq i$ then $X_{ip} \not\subseteq X_{ij}$, and if $p \neq h$ then $X_{hp} \not\subseteq X_{hj}$. By Lemma 4.2.4 it follows that since $X_{ip} \not\subseteq X_{ij}$ for all $p \in \{j + 1, \ldots, k \} \setminus \{i \}$, we have $X_{pj} = X_{pi}$ for all $p \in \{j + 1, \ldots, k \} \setminus \{i \}$. This implies that since $h > i$, we have $X_{hj} = X_{hi}$. Now there exists a $p \in \{j + 1, \ldots k \} \setminus \{h \}$ such that $X_{hj} = X_{hp}$; a contradiction to our assumption. This concludes the proof.

Lemma 4.2.6. Let $G$ be a perfect graph, let $J$ be 3-connected, and let $(S, P)$ be a good $J$-strip system. Let $X, Y \subseteq V(S, P)$ be serious $P$-sets. Then $X \cap Y \neq \emptyset$.

Proof. Let $v \in V(J)$. Since $J$ is 3-connected, it follows that $v$ has three distinct neighbors $w_1, w_2, w_3$ in $J$. Since $X, Y$ are serious $P$-sets, it follows that there exist $i, j \in \{1, 2, 3 \}$ such that $\bigcup_{k \in \{1, 2, 3 \} \setminus \{i \}} P_{vkw} \subseteq X$ and $\bigcup_{k \in \{1, 2, 3 \} \setminus \{j \}} P_{vkw} \subseteq Y$. Now let $k \in \{1, 2, 3 \} \setminus \{i, j \}$. It follows that $P_{vkw} \subseteq X \cap Y$, and thus $X \cap Y \neq \emptyset$.

Lemma 4.2.7. Let $G$ be a perfect graph, and let $J_1, \ldots, J_k$ be 3-connected graphs.

For $i \in \{1, \ldots, k \}$, we let $(S^i, P^i)$ be a good $J_i$-strip system in $G$. Suppose that $(S^1, P^1), \ldots, (S^k, P^k)$ is an increasing ordering of pairwise orthogonal strip systems. For $i \in \{1, \ldots, k \}$, we let $M^i$ denote the set of major vertices for $(S^i, P^i)$ in $G$.

For all $i, j \in \{1, \ldots, k \}$ with $i \neq j$, we define $X_{ij} = N^{pi}(V(M_{ij}))$, where $M_{ij}$ is an anticomponent of $M^i$ containing $M^i \cap V(S^j, P^j)$. For $i \in \{1, \ldots, k \}$, we let $Y_i$ be a serious $P^i$-set with $f_{(S^i, P^i)}(Y_i)$ maximum.
If $Z_i = \cap_{j \neq i} X_{ij} = \emptyset$, then there exist distinct $j, l \in \{1, \ldots, i - 1\}$ such that for all $p \in \{1, \ldots, k\} \setminus \{i\}$, if $p > j$ then $X_{ip} \not\subseteq X_{ij}$, and if $p > l$ then $X_{ip} \not\subseteq X_{il}$.

**Proof.** By Lemma 4.2.3 it follows that for all $p \in \{i + 1, \ldots, k\}$, we have $Y_i \subseteq X_{ip}$, and $Y_i$ is a serious $P^i$-set.

Suppose that $Z_i = \emptyset$. It follows that we may assume that there is a $j \in \{1, \ldots, i - 1\}$ such that $Y_i \not\subseteq X_{ij}$; choose $j$ maximum with this property.

Since $Y_i \subseteq X_{ip}$ for all $p \in \{j + 1, \ldots, k\} \setminus \{i\}$, it follows that $X_{ip} \not\subseteq X_{ij}$ for all $p \in \{j + 1, \ldots, k\} \setminus \{i\}$.

Now let $l \in \{1, \ldots, j - 1\}$. Suppose that for all $l \in \{1, \ldots, j - 1\}$, we have either $Y_i \subseteq X_{il}$ or $X_{ij} \subseteq X_{il}$. It follows that $X_{il} \cap X_{ij} \subseteq Z_i$. Since $X_{il}, X_{ij}$ are serious $P^i$-sets, it follows that $Z_i \neq \emptyset$ by Lemma 4.2.6, a contradiction. This proves that there exists an $l \in \{1, \ldots, j - 1\}$ such that $Y_i \not\subseteq X_{il}$ and $X_{ij} \not\subseteq X_{il}$. We choose $l$ maximum with this property. Since either $Y_i \subseteq X_{ip}$ or $X_{ij} \subseteq X_{ip}$ for all $p \in \{l + 1, \ldots, k\} \setminus \{i\}$, it follows that $X_{ip} \not\subseteq X_{il}$ for all $p \in \{l + 1, \ldots, k\} \setminus \{l\}$. Now $l$ and $j$ have the desired properties.

We are now ready to prove the main result of this chapter.

**Theorem 4.2.8.** Let $G$ be a perfect graph, and let $J_1, \ldots, J_k$ be 3-connected graphs. For $i \in \{1, \ldots, k\}$, we let $(S^i, P^i)$ be a good $J_i$-strip system in $G$. Suppose that $(S^1, M^1), \ldots, (S^k, M^k)$ are pairwise orthogonal. Then $k \leq 2\omega(G)$.

**Proof.** By symmetry, we may assume that $(S^1, P^1), \ldots, (S^k, P^k)$ is an increasing ordering. For $i \in \{1, \ldots, k\}$, we let $M^i$ denote the set of major vertices for $(S^i, P^i)$ in $G$.

For all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, we define $X_{ij} = N^{P^j}(V(M_{ij}))$, where $M_{ij}$ is an anticomponent of $M^i$ containing $M^i \cap V(S^j, P^j)$. For $i \in \{1, \ldots, k\}$, we let $Y_i$ be a serious $P^i$-set with $f_{(S^i, P^i)}(Y_i)$ maximum.
For \( i, j \in \{1, \ldots, k\} \), we call \((i, j)\) a bad pair if \( i < j \) and for all \( p \in \{j + 1, \ldots, k\} \setminus \{i\} \), \( X_{ip} \not\subseteq X_{ij} \). Let \( \mathcal{B} \) be the set of all bad pairs.

By Lemma 4.2.5 it follows that \( |\mathcal{B}| \leq k \). By Lemma 4.2.7 it follows that if \( Z_i = \emptyset \), then there exist distinct \( j, l \in \{1, \ldots, k\} \) such that \((j, i), (l, i) \in \mathcal{B}\).

It follows that \( |\{ i : Z_i = \emptyset \}| \leq |\mathcal{B}|/2 \), and thus \( |\{ i : Z_i \neq \emptyset \}| \geq k - |\mathcal{B}|/2 \). Since the sets \( Z_i \) for \( i \in \{1, \ldots, k\} \) are pairwise disjoint and complete to each other, it follows that \( G \) contains a clique of size at least \( k - |\mathcal{B}|/2 \). This implies that \( 2\omega(G) \geq 2k - |\mathcal{B}| \geq 2k - k = k \), as claimed. \( \square \)
Chapter 5

Conclusion

In this dissertation, we considered several problems related to cliques, stable sets, and coloring in graphs with forbidden induced subgraphs. Even though the three main chapters all deal with fairly different problems in this area, there are some interesting parallels between the Erdős-Hajnal and Gyárfás-Sumner conjectures, which we will discuss here. In particular, Conjecture 2.0.1 and Conjecture 3.2.3 have some similarities: both conjecture that a certain property holds if and only if $H$ is a forest, and for both of them, Erdős’ construction (see [20] and Theorems 2.0.2 and 3.2.5) is used to show that “only if” holds.

In fact, in many cases, we can state results from Chapter 3 in terms of graph coloring, as follows. Let $G$ be a graph. A function $\mu : 2^{V(G)} \to [0, \infty)$ is an outer-measure on $G$ if

- $\mu(\emptyset) = 0$;
- for all $A, B \subseteq V(G)$, $\mu(A) \leq \mu(A \cup B) \leq \mu(A) + \mu(B)$; and
- $\mu(V(G)) = 1$.

Let $\mu$ be an outer-measure on $G$. For $\varepsilon > 0$, we say that $(G, \mu)$ is $\varepsilon$-coherent if

- for all $v \in V(G)$, we have that $\mu(N[v]) \leq \varepsilon$; and
• for every anticomplete pair $A, B$, either $\mu(A) \leq \varepsilon$ or $\mu(B) \leq \varepsilon$.

For a graph $G$, if we let $\mu(A) = |A|/|V(G)|$ for all $A \subseteq V(G)$, then $G$ is $\varepsilon$-coherent if and only if $(G, \mu)$ is $\varepsilon$-coherent.

In [34], it was proved that for every caterpillar $T$, there exists an $\varepsilon > 0$ such that if $\mu$ is an outer-measure on $G$, and $(G, \mu)$ is $\varepsilon$-coherent, then $G$ contains $T$ as an induced subgraph. Our proof for Theorem 3.2.37 does not work for outer-measures as written; and so the following strengthening of Conjecture 3.2.3 is open:

**Conjecture 5.0.1.** For every tree $T$ there exists an $\varepsilon > 0$ such that if $\mu$ is an outer-measure on $G$, and $(G, \mu)$ is $\varepsilon$-coherent, then $G$ contains $T$ as an induced subgraph.

Results we prove for outer-measures have consequences for coloring; for a graph $G$, we define the outer-measure $\mu_\chi$ by setting $\mu_\chi(A) = \chi(A)/\chi(G)$ for all $A \subseteq V(G)$.

Proofs related to the Gyárfás-Sumner conjecture are usually by induction on clique number, and since for every graph $G$, $\omega(N(v)) < \omega(G)$ for all $v \in V(G)$, we can generally assume that $\mu_\chi(N[v]) \leq \varepsilon$. Now, if $(G, \mu)$ is not $\varepsilon$-coherent, it follows that $G$ contains an anticomplete pair $A, B$ with $\chi(A), \chi(B) \geq \varepsilon \chi(G)$. While this is not as useful for Conjecture 2.0.1 as having an anticomplete $(\varepsilon n, \varepsilon n)$-pair is for Conjecture 3.1.1 it is still surprising. For further details, see [16].

Despite these similarities, the results in Table 5.1 show that if a class of graphs is $\chi$-bounded, it does not follow that there exists an $\varepsilon > 0$ such that no graph in the class in $\varepsilon$-coherent; and conversely, if there exists an $\varepsilon > 0$ such that no graph in the class is $\varepsilon$-coherent, that does not imply that the class is $\chi$-bounded.
<table>
<thead>
<tr>
<th>Class</th>
<th>$\chi$-bounded</th>
<th>not $\varepsilon$-coherent</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_k$-free</td>
<td>yes 29</td>
<td>yes 4</td>
</tr>
<tr>
<td>$T$-free, $T$ a subdivided caterpillar</td>
<td>open</td>
<td>yes, Theorem 3.2.37</td>
</tr>
<tr>
<td>$T$-free, $T$ a tree of radius 2</td>
<td>yes, by definition 15</td>
<td>no 25</td>
</tr>
<tr>
<td>perfect</td>
<td>open, Conjecture 2.0.1</td>
<td>no 25</td>
</tr>
<tr>
<td>no long odd hole</td>
<td>yes 13</td>
<td>yes, Theorem 3.2.37</td>
</tr>
<tr>
<td>$T$-free, $T$ a tree</td>
<td>no induced subdivision of $H$</td>
<td>yes, Theorem 3.2.18</td>
</tr>
<tr>
<td>no directed star</td>
<td>no for all $H$ [6]</td>
<td>no, Theorem 3.2.43</td>
</tr>
</tbody>
</table>

Table 5.1: Results for $\chi$-boundedness and not being $\varepsilon$-coherent for different graph classes. Here, “not $\varepsilon$-coherent” means that there exists an $\varepsilon > 0$ such that no graph in the class in question is $\varepsilon$-coherent.
Bibliography


