Properties of Typical T-Free Graphs

Yelena Yuditsky

SCHOOL OF COMPUTER SCIENCE

MCGILL UNIVERSITY, MONTREAL

DECEMBER, 2017

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Dedication

To the memory of my parents, Mera and Yacov Yuditsky.

Contents

List of Figures

Abstract

In this thesis we study the structure of almost all T-free graphs for any tree T, that is graphs that do not contain T as an induced subgraph. Let $\nu(T)$ be the size of a maximum matching in T. We show that almost all T-free graphs can be partitioned into $w(T) := |V(T)| - \nu(T) - 1$ parts such that each part has a specific structure. Specifically, each part is F-free for all $F \in \mathcal{F}$, where $\mathcal{F}$ is a collection of graphs on at most 4 vertices. This result allows us to show that for every tree T, for almost all T-free graphs G, every induced subgraph G' of G satisfies $\chi(G') \leq w(T) \cdot \omega(G')$. Moreover, for every tree T, almost all T-free graphs G have $\chi(G) = \omega(G)$. This proves an asymptotic version of the Gyárfás-Sumner conjecture.

Abrégé

Dans cette thèse on s'intéresse à la structure typique d'un graphe sans sous-graphe induit isomorphe à un certain arbre T donné. Soit $\nu(T)$ la taille d'un couplage maximum dans T, et soit $w(T) := |V(T)| - \nu(T) - 1$. On montre que presque tous les arbres sans T induit peuvent être partitionnés en $w(T)$ parties, toutes d'une structure très particulière. Plus précisément, chaque partie est sans F induit, pour tout $F \in \mathcal{F}$ où $\mathcal{F}$ est une collection de graphes à au plus 4 sommets. Ce résultat nous permet de prouver que pour tout arbre T, presque tout graphe G sans F induit est tel que tout sous-graphe induit H de G satisfait $\chi(H) \leq w(T) \cdot \omega(H)$. De plus, pour tout arbre T, presque tout graphe G sans T induit satisfait $\chi(G) = \omega(G)$. Ceci confirme une version asymptotique de la conjecture de Gyárfás-Sumner.

Acknowledgements

I would like to thank my supervisor Bruce Reed. Without Bruce this thesis would not have been possible. I learned a lot from working with Bruce and I am thankful to him for that. I would also like to thank Bruce for his generous support over the years. Due to this support I had the opportunity to attend many conferences and summer schools, allowing me to stay informed of the latest results, developments and techniques in my field of research. It also gave me an opportunity to meet many great mathematicians and engage in collaborations with them. I would also like to thank Bruce for the numerous hours he spent in trying to improve my writing.

I would like to thank my co-suprevisor Sergey Norin. Sergey's aid and guidance towards the end of the writing of my thesis was invaluable to me. I would like to thank Sergey for his patience and his kindness. I deeply appreciate all that he has done for me.

I would like to thank Bruce, Sergey, Louigi Addario-Berry, Hamed Hatami and Bruce Shepherd for the exciting courses that I was privileged to take with them. These courses expanded my horizons and exposed me to diverse fields of mathematics. They helped me to grow as a mathematician.

I would like to thank my friends Marthe Bonamy and Tash Morrison. My long conversations with Marthe have always been a pleasure. I would like to thank her constant willingness to help and listen even in her extremely tight schedule. I would like to thank Tash for her support and encouragement when I needed it. I thank her for the many enjoyable walks we took together during our travels in France, Germany, Switzerland and Brazil. I also want to thank Marthe and Tash for their many *hug*s.

During my studies at McGill I was fortunate enough to meet many amazing people. I would like to thank Xing Shi Cai, Laura Eslava, Lianna Hambardzumyan, Yaqiao Li, Atefeh Mohajeri, Jon Noel, Ying Jie Qian and Liana Yepremyan. It was always a joy to take courses and to work on assignments together. Due to them, assignments became something to look forward to with anticipation. I would like to thank them for all the lunches, dinners and coffee breaks we had together. It was always a delight.

I would like to thank to my friends Nova Fandina, Miri Gubenko, Ilan Karpas, Katia Kristalny and Igal Shilman. They always kept in touch with me although I was an ocean apart and they always made me feel at home every time I came to visit.

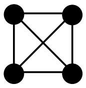
Finally, I would like to thank to Nima Hoda for his endless support and help. Without his assistance the last few years would have been so much more difficult. Also I would like to thank Nima's family, Ezzat, Hashem, Shiva, Reza, Misha and Arteen, for becoming as my own.

Preface

The main contributions of this thesis are contained in Chapters 3, 4, 5 and 6. All the results are joint work with Prof. Bruce Reed.

Chapter 1

Introduction

Graphs are very natural and concrete objects which arise frequently in daily life. For example, while browsing the internet, taking the metro to work or trying to challenge one nephew by asking him to draw  without lifting his pencil from the paper. Graphs also appear in many branches of science, for example, in computer science, biology, physics, linguistics, sociology and elsewhere. Graphs model a symmetric relation between entities. We call those entities **vertices**. If there is a relation between two vertices, then we say that there is an **edge** between those two vertices.

The study of graphs without some fixed substructure have received considerable attention during the last century. There are two ways in which substructures are commonly forbidden in graphs. A graph H with h vertices is a **weak subgraph** of a graph G, if it is possible to find in G a set X of h vertices where by taking a subset of the edges in G all of whose ends lie in X, we get a copy of H on X. For example, a graph G with n vertices and all the possible edges between the vertices contains any graph H on at most n vertices as a weak subgraph. A graph H with h vertices is an **induced subgraph** of G if it is possible to find in G a set X of h vertices where by taking *all* the edges in G with both ends in X, we get a copy of H on X. For example, a graph G with n vertices and all the possible edges between the vertices contains as an induced subgraphs only graphs H on at most n vertices and all the possible edges between the vertices of H. Let H be a graph. A graph G is ***H*-free** if it does not contain an *induced* copy of H.

There are many interesting questions regarding the properties of graphs which do not contain some graph (or graphs) as an induced subgraph. We mention some of the most famous such questions. We start with the Erdős-Hajnal conjecture [25]. Roughly speaking, the conjecture asserts that for any graph H, any graph G which is H-free contains a large set of vertices X such that either there are all the possible edges between any pair of vertices on X (the graph induced on X is a **clique**) or there are no edges on X at all (the graph induced on X is a **stable set**). Large here means that the set X must be of at size at least $|V(G)|^{\varepsilon}$ for some $\varepsilon > 0$. For a graph G, we denote by $\boldsymbol{V(G)}$ the vertex set of G and by $\boldsymbol{E(G)}$ the edge set of G.

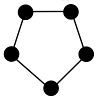
Conjecture 1.0.1 (Erdős-Hajnal [25])**.** *Let H be a graph, then there is an $\varepsilon = \varepsilon(H) > 0$, such that every H-free graph contains a stable set or a clique of size at least $|V(G)|^{\varepsilon(H)}$.*

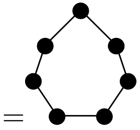
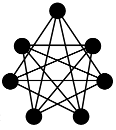
The Erdős-Hajnal conjecture is still open and appears to be very difficult to resolve (see a survey in [17]).

In some cases it is useful to partition a given graph G into a minimal number of stable sets. For example, in the famous four colour problem, we want to colour countries on a map with only 4 colours, such that any two neighbouring countries receive different colours. For this problem it is possible to define a graph whose vertex set are the different countries, and two vertices which represent two countries have an edge between them if and only if they are neighbouring on the map. A stable set in this case represent a set of countries such that no two countries share a border. The number of colours needed to colour the map is equal to the minimal number of stable sets in which we can partition the graph. The **chromatic number** of a graph G, which is denoted as $\boldsymbol{\chi(G)}$, is the minimal number of stable sets into which it is possible to partition the vertex set $V(G)$ of G. Equivalently, $\chi(G)$ is the minimal number k such that there is a function $c : V(G) \to [k]$ where for each $\{v_1, v_2\} \in E(G)$, $c(v_1) \neq c(v_2)$. The graph which we get from the neighbour relation of the countries in a map is an example of a **planar graph**. A planar graph is a graph which can be drawn in the plane such that any two edges can intersect only at a vertex. The following result was shown by Appel and Haken in 1976.

Theorem 1.0.2 (The Four Colour Theorem [3])**.** *Let G be a planar graph, then $\chi(G) \leq 4$.*

An immediate observation that one can make is that the chromatic number of a graph G must be at least the size of a maximum clique in G. This is true because no two vertices in a clique can be in the same stable set (equivalently, receive the same colour). We call the size of a maximum clique G, the **clique number** of G, and we denote it by $\boldsymbol{\omega(G)}$.

The reverse inequality $\chi(G) \leq \omega(G)$ does not hold in general. A **path** in a graph G is a sequence of vertices $v_1, v_2, ...v_k$ such that all the vertices are different, and for every $1 \leq i \leq k-1$, $\{v_i, v_{i+1}\} \in E(G)$. A **cycle** in a graph G is a path where the first and last vertices are the same. A simple example of a graph G for which $\chi(G) \leq \omega(G)$ is false is a cycle on 5 vertices $C_5 =$ . For C_5, $\omega(C_5) = 2$ and $\chi(C_5) = 3$.

A graph G in which every induced subgraph G' satisfies, $\chi(G') \leq \omega(G')$, is called a **perfect graph**. Claude Berge conjectured in 1963 [11], that a graph is perfect if and only if it does not contain as an induced subgraph an odd cycle with at least 5 vertices or the complement of an odd cycle with at least 5 vertices, where the **complement** of a graph G is the graph on the same vertex set but with the set of edges being the complement of the set of edges in G. For example the complement of $C_7 =$  is $\overline{C_7} =$ . Berge's conjecture was open for 40 years and was proved in the groundbreaking work by Chudnovsky, Robertson, Seymour and Thomas [18].

Theorem 1.0.3 (The Strong Perfect Graph Theorem [18])**.** *A graph G is a perfect graph if and only if no induced subgraph of G is an odd cycle with at least 5 vertices or a complement of one.*

As was mentioned it is not true that $\chi(G) \leq \omega(G)$ for any G, but one can ask if it might be the case that for any G, $\chi(G)$ can be bounded by some larger function of $\omega(G)$, as, for example, by $\omega(G)^2$ or $2^{\omega(G)}$. The answer for that question is also no. Mycielski [40] gave

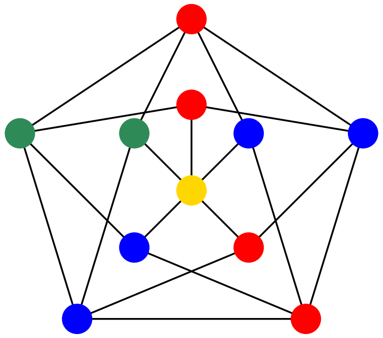


Figure 1.1: G' constructed from C_5.

a construction of a family of graphs with clique number 2 and arbitrarily large chromatic number. We call a clique on 3 vertices a **triangle**.

Theorem 1.0.4 (Mycielski [40]). *Let G be a triangle-free graph, and let G' be a graph constructed from G as follows,*

- *The vertex set of G' is $V \cup V' \cup \{u\}$ where V and V' are two copies of the vertex set of G.*
- *The graph induced on V is G.*
- *Let $v' \in V'$ be a vertex which is a copy of the vertex v in V, then the vertices to which v' is adjacent are exactly the vertices to which v is adjacent.*
- *The vertex u is adjacent to all the vertices in V'.*

Then the graph G' is triangle-free and $\chi(G') = \chi(G) + 1$.

The graphs in Mycielski's construction are triangle-free, but might contain cycles on 4 vertices. One can ask further if it might be the case that if a graph G does not contain short cycles then it is possible to bound $\chi(G)$ with a function of $\omega(G)$. The answer to this question is also no. Erdős showed that for any $g \in \mathbb{N}$, there are graphs with no induced cycles with less than g vertices and arbitrarily large chromatic number [21]. The **length of a path** (or the **length of a cycle**) is the number of vertices in it. The **girth** of a graph G is the length of the shortest induced cycle in G.

Theorem 1.0.5 (Erdős [21])**.** *Let* $g, c \in \mathbb{N}$*. Then there exists a graph* G *with girth larger than* g *and chromatic number larger than* c*.*

As it was mentioned it is not true that the chromatic number of any graph can be bounded by its clique number (or its girth), therefore we restrict ourselves to smaller families of graphs. A **hereditary family** (or **hereditary property**) of graphs is a family which is closed under taking induced subgraphs. A hereditary family $\mathcal{F}$ is **χ-bounded**, if there is a bounding function $f_{\mathcal{F}}$, such that for each $G \in \mathcal{F}$, $\chi(G) \leq f_{\mathcal{F}}(\omega(G))$. By Theorems 1.0.4 and 1.0.5 the family of all graphs is not χ-bounded. By Theorem 1.0.3 the family of graphs which do not contain an induced cycle of length at least 5 or a complement of one are perfect and therefore χ-bounded with the bounding function being the identity. We restrict ourselves to families which are defined by forbidding some graphs as an induced subgraphs.

Let H be a graph which contains a cycle of length ℓ as an induced subgraph. By Theorem 1.0.5, there are graphs with girth greater than ℓ, and therefore which are H-free, and which have arbitrarily large chromatic number. Therefore the family of H-free graphs for a graph H which contains a cycle is not χ-bounded. Now let H be a graph without induced cycles. We call such a graph H a **forest**. Gyárfás [30] and Sumner [54] independently conjectured that families of graphs without induced forest H are χ-bounded. We denote by **Forb(H)** the family of all H-free graphs.

Conjecture 1.0.6 (Gyárfás-Sumner conjecture [30], [54])**.** *Let* H *be a forest and let* $\mathcal{F} = \text{Forb}(H)$ *be the family of all* H*-free graphs, then there is a function* $f_{\mathcal{F}}$ *such for any graph* $G \in \mathcal{F}$*,* $\chi(G) \leq f_{\mathcal{F}}(\omega(G))$*.*

A **tree** is a connected forest. A **connected graph** is a graph where there is a path between any two vertices. Note that the family Forb(H), for a forest H, is χ-bounded if and only if all the families Forb(T) are χ-bounded for each of the connected components T of H. Hence the above conjecture can be reduced to trees. The Gyárfás-Sumner conjecture has been proved for the following families of trees which we mostly define later: paths and stars [30], trees of radius two [35], trees which are subdivided stars [50], trees obtained from trees of radius two by making exactly one subdivision in every edge adjacent to the root [36], "two-legged caterpillars", "double-ended brooms" and a few others [19]. The conjecture is still open in its general form.

The above mentioned results and conjectures ask about the properties of *all* graphs in some family of graphs. It is of interest to study the properties of *almost all* graphs in a family. Let $\mathcal{F}$ be a family of graphs and let $n \in \mathbb{N}$. We denote by $\boldsymbol{(\mathcal{F})_n} \subseteq \mathcal{F}$ the set of graphs in $\mathcal{F}$ on exactly n vertices. For two families $\mathcal{F}, \mathcal{F}'$, such that $\mathcal{F} \subseteq \mathcal{F}'$, we say that **almost all** graphs in $\mathcal{F}'$ are in $\mathcal{F}$ if

$$\lim_{n\to\infty} \frac{|(\mathcal{F})_n|}{|(\mathcal{F}')_n|} = 1.$$

Sometimes we refer to the graphs in the above family $\mathcal{F}$ as **typical** graphs in $\mathcal{F}'$. The Erdős-Hajnal conjecture was shown to be true for almost all graphs without some fixed graph H as an induced subgraph [39].

Theorem 1.0.7 ([39])**.** *Let H be a graph, then there exists an $\varepsilon = \varepsilon(H) > 0$ such that almost all H-free graphs G contain a stable set or a clique of size at least $|V(G)|^{\varepsilon}$.*

The above theorem was strengthened in the case where we do not restrict ourselves to all graphs H [33].

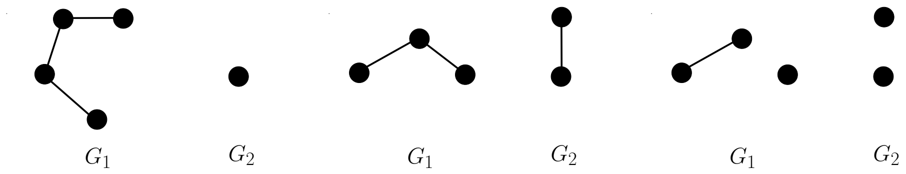
Theorem 1.0.8 ([33])**.** *For almost all graphs H, there is a constant $b = b(H) > 0$ such that almost all H-free graphs G contain a stable set or a clique of size at least $b|V(G)|$.*

The authors in [33] conjectured that the above theorem might be true for all graphs H besides P_3 and P_4 which are paths on 3 and 4 vertices, respectively.

Before the strong perfect graph conjecture was proved, Prömel and Steger [45] showed that almost all graphs without an induced C_5 are perfect.

In this thesis, we show that the Gyárfás-Sumner conjecture holds for almost all T-free graphs for any tree T.

Theorem 1.0.9. *For every tree T, almost all T-free graphs G have $\chi(G') \leq \big(|V(T)| - \nu(T) - 1\big) \cdot \omega(G')$ for every induced subgraph G' of G. Moreover, for every tree T, almost all T-free graphs G have $\chi(G) = \omega(G)$.*



(a) P_4 and a vertex (b) P_3 and K_2 (c) $\overline{P_3}$ and S_2

Figure 1.2: Partitions of C_5.

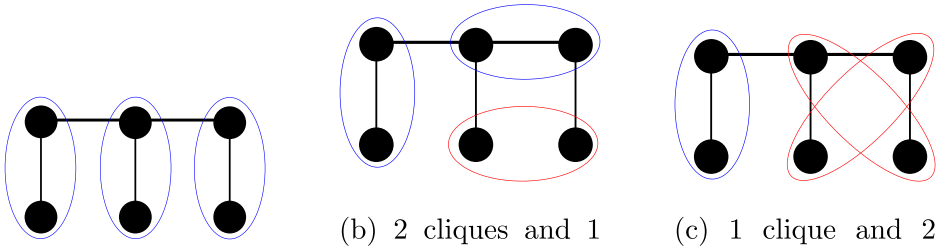
We prove this theorem by first obtaining strong structural results for almost all graphs without some fixed induced tree T. Then we use the structure to show that almost all T-free graphs have the required colouring.

1.1 Our Results and Outline of the Thesis

In this subsection we present our structural results. Before we can state those results we present a few additional definitions.

Let G be a graph and let $V' \subseteq V(G)$, we denote by $\boldsymbol{G[V']}$ the subgraph of G induced on V'. Let G be a graph and let $V' \subseteq V(G)$, we denote by $\boldsymbol{G \setminus V'}$ the induced subgraph $G[V(G) \setminus V']$. Let G' be a subgraph of G, we denote by $\boldsymbol{G \setminus G'}$ the induced subgraph $G \setminus V(G')$. A **partition** of a graph G is a collection $\{G[V_1], G[V_2], ..., G[V_k]\}$ where $\{V_1, V_2, ..., V_k\}$ is a partition of $V(G)$. See for example in Figure 1.2 some possible partitions of C_5.

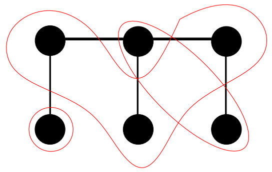
Let H be a graph and let G be a graph which has a partition into $\chi(H)-1$ stable sets, then such a graph G is H-free. This is true because otherwise, a partition of G into $\chi(H)-1$ stable sets would induce a partition of $V(H)$ into $\chi(H)-1$ stable sets which is a contradiction to the definition of the chromatic number $\chi(H)$ of H. Using the same idea, if G has a partition into $\sigma(H)-1$ cliques, where $\boldsymbol{\sigma(H)}$ is minimal number of cliques into which H can be partitioned, then such a graph G is H-free. We do not have to restrict ourselves to partitions of G into only stable sets or into only cliques. Let G be a graph which can be partitioned into s stable sets and c cliques for values s, c such that H cannot be partitioned into s stable sets and c cliques, then again such a graph G is H-free. For any such s, c we can consider the set of all



(a) 3 cliques

(b) 2 cliques and 1 stable sets

(c) 1 clique and 2 stable sets



(d) 3 stable sets (P_3^* can be partitioned into 2 stable sets)

Figure 1.3: Partitions of P_3^* into cliques and stable sets.

graph which can be partitioned into s stable sets and c cliques. We get the biggest set of such graphs G when the values s, c are chosen so $s + c$ is as large as possible (and H cannot be partitioned into s stable sets and c cliques). The **witnessing partition number** of a graph H, in short **wpn($\boldsymbol{H}$)**, is equal to the maximal sum $s + c$ such that there are s, c and H cannot be partitioned into s stable sets and c cliques.

A **leaf** of a tree is a vertex which has exactly one neighbour. Consider the graph which is a path on 3 vertices with one leaf attached to each of the vertices on the path, we denote this graph as P_3^*. The graph P_3^* cannot be partitioned into two cliques, but it can be partitioned into 3 cliques, into two stable sets and one clique, into two cliques and one stable set, and into 3 stable sets, see Figure 1.3 for partitions of P_3^*, therefore $\text{wpn}(P_3^*) = 2$.

Prömel and Steger [46] were the first to introduce the above idea of the witnessing partition number of a graph. They considered the value which they denoted by $\tau(H)$ and is equal to $\text{wpn}(H) + 1$. In the literature this number also sometimes referred as binary chromatic (in short, bichromaric) number (e.g. [4]) or a colouring number of a graph (e.g. [12]).

There are many classical results, which are stated as a function of the chromatic number of H, regarding families of graphs which do not contain some graph H as a weak subgraph. Prömel and Steger [44, 46] showed that similar results remain true also for families of H-free graphs if the chromatic number is exchanged by $\tau(H) = \mathrm{wpn}(H) + 1$.

As described earlier, for a given graph H, every graph G which can be partitioned into s stable sets and c cliques such that $s + c = \mathrm{wpn}(H)$ and H does not have such a partition, is H-free. This is not the only way in which we can obtain an H-free graph. Let $P = (H_1, H_2, ..., H_w)$ be some ordered partition of a graph H. We say that an ordered partition $(G_1, G_2, ..., G_w)$ of a graph G is $\boldsymbol{P}$**-free** if there is an $i \in [w]$ such that G_i is H_i-free. Let $\boldsymbol{P(H)}$ be a set of all ordered partitions $(H_1, H_2, ..., H_{\mathrm{wpn}(H)})$ of H into some sequence of $\mathrm{wpn}(H)$ subgraphs. Let G be a graph which has an ordered partition $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$ which is P-free for any partition $P \in P(H)$, then G is H-free. We call such a partition $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$, a $\boldsymbol{P(H)}$**-free** partition. Reed and Scott made the following conjecture.

Conjecture 1.1.1 (Reed-Scott [47]). *For every graph H, almost all H-free graphs G have a $P(H)$-free partition.*

There are a few graphs for which the above conjecture is verified. It is known for graphs H which are cliques [38], and a special set of graphs which are called *critical graphs* [5]. Conjecture 1.1.1 is also proved for cycles of odd length [44, 5], and for cycles of even length [43, 47, 34].

The main result of this thesis is a proof of Conjecture 1.1.1 in the case where H is a tree. Let us now present some of the additional necessary definitions. A **matching** in G is a collection of disjoint edges from $E(G)$. Let $\boldsymbol{\nu(T)}$ be the size of the largest matching in T. We say that a tree T has a **perfect matching** if $2\nu(T) = |V(T)|$.

A disjoint union of two disjoint graphs G_1 and G_2 is a new graph $\big(V(G_1) \cup V(G_2), E(G_1) \cup E(G_2)\big)$. A **join** of two disjoint graphs G_1 and G_2 is a new graph $\big(V(G_1) \cup V(G_2), E(G_1) \cup E(G_2) \cup E'\big)$ where E' is the set of all edges with one end in $V(G_1)$ and the other in $V(G_2)$. A **star** is the graph which is a complement of the disjoint union of a vertex and a clique. We say that a graph $G = (V, E)$ is **bipartite** if we can partition G into two disjoint stable sets A, B. We say that a graph $G = (V, E)$ is ***r*-partite** if we can partition G into r disjoint stable sets $V_1, V_2, ..., V_r$.

Let H be a graph, we call a sequence of families $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ a $\boldsymbol{P(H)}$**−free sequence** if for any choice of graphs $G_i \in \mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, the resulting partition $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$ is $P(H)$-free. Moreover, the families are maximal and hereditary, and to make the presentation easier, we restrict the families to contain graphs with at least h vertices where $|V(H)| = h$. Note that there might be more than one $P(H)$-free sequence for a graph H.

To state our structural results we need to partition the set of trees into the following families. Trees T with $2\nu(T) \leq |V(T)| - 2$, trees T with $2\nu(T) = |V(T)| - 1$ and trees T with $2\nu(T) = |V(T)|$, that is trees with a perfect matching. We partition further the set of trees with a perfect matching into families which are defined below. The reason for the above partition into the different families is that the trees in each family have different possible partitions $P(T)$ into smaller graphs. In particular, this implies that the resulting $P(T)$-free sequences are different for the trees in the different families. For each of the families we characterize the $P(T)$-free sequences for the trees T in the corresponding family. Then we state the theorem regarding the typical structure of T-free graphs for trees in each family.

Theorem 1.1.2 ($\mathcal{P}(T)$-free sequence for trees T with $2\nu(T) \leq |V(T)| - 2$)**.** *Let T be a tree such that $2\nu(T) \leq |V(T)| - 2$, and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be a $\mathcal{P}(T)$-free sequence, then for every $i \in [\mathrm{wpn}(T)]$, $\mathcal{F}_i$ is the set of all cliques.*

Theorem 1.1.3 (Structure of typical T-free graph for trees T with $2\nu(T) \leq |V(T)| - 2$)**.** *Let T be a tree with $2\nu(T) \leq |V(T)| - 2$, then almost all T-free graphs can be partitioned into $\mathrm{wpn}(T)$ parts such that each part is a clique.*

A **subdivided star** is a graph which is a star where each edge is subdivided exactly once.

Theorem 1.1.4 ($\mathcal{P}(T)$-free sequence for trees T with $2\nu(T) = |V(T)| - 1$ and T is a subdivided star)**.** *Let T be a tree with $|V(T)| \geq 5$ such that $2\nu(T) = |V(T)| - 1$ and T is a subdivided star.*

Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ *be the* $\mathcal{P}(T)$*-free sequence, then we have one of the following cases,*

(i) $\mathcal{F}_1$ *is the set of all stable sets, and* $\mathcal{F}_i$*, for* $2 \leq i \leq \text{wpn}(T)$*, is the set of all cliques.*

(ii) For every $i \in [\text{wpn}(T)]$*,* $\mathcal{F}_i$ *is the set of all cliques.*

Theorem 1.1.5 ($\mathcal{P}(T)$-free sequence for trees T with $2\nu(T) = |V(T)| - 1$ and T is not a subdivided star)**.** *Let* T *be a tree with* $|V(T)| \geq 5$ *such that* $2\nu(T) = |V(T)| - 1$ *and* T *is a subdivided star. Let* $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ *be the* $\mathcal{P}(T)$*-free sequence, then for every* $i \in [\text{wpn}(T)]$*,* $\mathcal{F}_i$ *is the set of all cliques.*

Theorem 1.1.6 (Structure of typical T-free graphs for T with $2\nu(T) = |V(T)| - 1$)**.** *Let* T *be a tree with* $2\nu(T) = |V(T)| - 1$ *and* $|V(T)| \geq 5$*. If* T *is a subdivided star, then almost all* T*-free graphs can be partitioned into* $\text{wpn}(T)$ *parts such that each part is a clique, or into* $\text{wpn}(T)$ *parts such that one of the parts is a stable set and the rest are cliques. If* T *is not a subdivided star then almost all* T*-free graphs can be partitioned into* $\text{wpn}(T)$ *parts such that each part is a clique.*

Note that Theorems 1.1.3 and 1.1.6 can be also derived from a result by Balogh and Butterfield [5]. In Section 5.1, we reprove their result.

The next results are for trees with a perfect matching, that is for trees T with $2\nu(T) = |V(T)|$. To state the results we define a few families of graphs. As mentioned $\boldsymbol{P_i}$ is a path on i vertices, $\boldsymbol{K_i}$ is a clique with i vertices and $\boldsymbol{S_i}$ is a stable set on i vertices for any $i \in \mathbb{N}$, note that $K_1 = S_1$. We denote by $G_1 + G_2$ the disjoint union of graph G_1 and G_2 and $G_1 + G_1 = 2G_1$. Let $\mathcal{H}$ be a collection of graphs, by **Forb($\mathcal{H}$)** we denote the set of all graphs which are H-free for each $H \in \mathcal{H}$. As mentioned earlier, in the case that $\mathcal{H} = \{H\}$ for some H then we write Forb(H) instead of Forb($\{H\}$).

- Let $\boldsymbol{\mathcal{G}_1}$ be Forb($\mathcal{H}$) for $\mathcal{H} = \{S_3, \overline{P_3}\}$.
- Let $\boldsymbol{\mathcal{G}_2}$ be Forb($\mathcal{H}$) for $\mathcal{H} = \{P_4, 2K_2, K_2 + S_2\}$.
- Let $\boldsymbol{\mathcal{G}_3}$ be Forb($\mathcal{H}$) for $\mathcal{H} = \{P_4, 2K_2, K_2 + S_2, S_4\}$.

- Let $\boldsymbol{\mathcal{G}_4}$ be Forb($\mathcal{H}$) for $\mathcal{H} = \{P_4, P_3 + K_1, 2K_2, K_2 + S_2\}$.
- Let $\boldsymbol{\mathcal{G}_5}$ be Forb($\mathcal{H}$) for $\mathcal{H} = \{P_4, P_3 + K_1, 2K_2, K_2 + S_2, S_4\}$.
- Let $\boldsymbol{\mathcal{G}_6}$ be Forb($\mathcal{H}$) for $\mathcal{H} = \{P_4, 2K_2, P_3 + S_1\}$.

It is not hard to see and we show in Section 5.2.1 the following.

Claim 1.1.7. *The families $\mathcal{G}_i$, $i \in [5]$ are as following.*

- *$\mathcal{G}_1$ is the family of graphs which are the join of stable sets of size at most 2 (multi-partite graph with parts of size at most 2).*
- *$\mathcal{G}_2$ is the family of graphs which are the join of complete multi-partite graph with an isolated vertex.*
- *$\mathcal{G}_3$ is the family of graphs which are the join of complete multi-partite graph with parts of size at most 2 and an isolated vertex.*
- *$\mathcal{G}_4$ is the family of graphs which are the join of graphs which are either stable sets or disjoint union of a vertex and a clique.*
- *$\mathcal{G}_5$ is the family of graphs which are the join of graphs which are either stable sets of size three or a disjoint union of a vertex and a clique.*
- *$\mathcal{G}_6$ is the family of graphs which are the join of graphs which are disjoin union of a stable set and a clique.*

We denote by $\boldsymbol{\mathcal{T}^{\mathrm{pl}}}$ be the set of trees with a perfect matching and where every non-leaf vertex has a neighbour which is a leaf. The set $\boldsymbol{\mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}} \subset \mathcal{T}^{\mathrm{pl}}$ is the family of all trees obtained from stars by subdividing every edge, except one, exactly once.

Theorem 1.1.8 ($\mathcal{P}(T)$-free sequence for trees $T \in \mathcal{T}^{\mathrm{pl}}$)**.** *Let T be a tree such that $T \in \mathcal{T}^{\mathrm{pl}}$ and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then we have one of the following cases,*

(i) $\mathcal{F}_i = \mathcal{G}_1$ for each $i \in [\mathrm{wpn}(T)]$.

(ii) The families can be reindexed such that $\mathcal{F}_1$ is the set $\mathcal{G}_2$ and the rest of the families are the sets of all cliques.

Theorem 1.1.9 (Structure of typical T-free graphs for $T \in \mathcal{T}^{\mathrm{pl}}$)**.** *Let $T \in \mathcal{T}^{\mathrm{pl}}$, then almost all T-free graphs can be partitioned into* $\mathrm{wpn}(T)$ *parts such that each part is a complete multi-partite graph with parts of size at most two.*

Theorem 1.1.10 ($\mathcal{P}(T)$-free sequence for trees $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$)**.** *Let T be a tree such that $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then we have one of the following cases,*

(i) The families can be reindexed such that $\mathcal{F}_1, \mathcal{F}_2$ are the set $\mathcal{G}_1$, and the rest of the families are the sets of all the cliques.

(ii) The families can be reindexed such that $\mathcal{F}_1$ is the set $\mathcal{G}_3$ and the rest of the families are the sets of all the cliques.

Theorem 1.1.11 (Structure of typical T-free graphs for $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$)**.** *Let $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, then almost all T-free graphs can be partitioned into* $\mathrm{wpn}(T)$ *parts such that two of the parts are complete multi-partite graphs with parts of size at most two, and the rest are cliques.*

The set of trees which have a perfect matching but are not in $\mathcal{T}^{\mathrm{pl}}$, we denote by $\boldsymbol{\mathcal{T}^{\mathrm{npl}}}$. A tree T is in $\boldsymbol{\mathcal{S}} \subset \mathcal{T}^{\mathrm{npl}}$ if there is some path P in T of length 6 or 8, such that the ends of P are leaves and the following hold. Let $\mathcal{C}$ be the set of connected components in $T \setminus P$, then each components in $\mathcal{C}$ is an edge with the following additional properties.

i. For $P = v_1, v_2, v_3, v_4, v_5, v_6$ of length 6, each of the components in $\mathcal{C}$ is joined by an edge to P at either v_3 or v_4. We denote this set as $\mathcal{S}^*_6$.

ii. For $P = v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8$ of length 8, each of the components in $\mathcal{C}$ is joined by an edge to P at either v_3 or v_6. We denote this set as $\mathcal{S}^*_8$.

Let $\mathcal{S} = \mathcal{S}^*_6 \cup \mathcal{S}^*_8$.

Theorem 1.1.12 ($\mathcal{P}(T)$-free sequence for trees $T \in \mathcal{T}^{\text{npl}} \setminus \{P_6\}$)**.** *Let* $T \in \mathcal{T}^{\text{npl}} \setminus \{P_6\}$*, and let* $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ *be the* $\mathcal{P}(T)$*-free sequence, then the families can be reindexed such that* $\mathcal{F}_1$ *is* $\mathcal{G}_4$*, and the rest of the families are the sets of all cliques.*

Theorem 1.1.13 (Structure of typical T-free graphs for $T \in \mathcal{T}^{\text{npl}} \setminus \{P_6\}$)**.** *Let* $T \in \mathcal{T}^{\text{npl}} \setminus \{P_6\}$*, then almost all* T*-free graphs can be partitioned into* wpn(T) *parts such that one of the parts is a complement of a disjoint unions of cliques and stars, and the rest are cliques.*

Theorem 1.1.14 ($\mathcal{P}(T)$-free sequence for the tree $T = P_6$)**.** *Let* $T = P_6$ *and let* $(\mathcal{F}_1, \mathcal{F}_2)$ *be the* $P(P_6)$*-free sequence, then we have the following cases,*

(i) The families can be reindexed such that $\mathcal{F}_1$ *is the family of all stable sets and* $\mathcal{F}_2 = \mathcal{G}_1$*.*

(ii) The families can be reindexed such that $\mathcal{F}_1$ *is the family of all cliques and* $\mathcal{F}_2 = \mathcal{G}_6$*.*

Theorem 1.1.15 (Structure of typical T-free graphs for $T = P_6$)**.** *Let* $T = P_6$*, then almost all* T*-free graphs can be partitioned into* wpn$(P_6) = 2$ *parts such that one of the parts is a clique and the other is a join of graphs which are disjoint union of a stable set and a clique.*

Theorem 1.1.16 ($\mathcal{P}(T)$-free sequence for trees $T \in \mathcal{T}^{\text{npl}} \setminus (\mathcal{S} \cup \{P_6\})$)**.** *For a tree* $T \neq P_6$ *such that* $T \in \mathcal{T}^{\text{npl}} \setminus (\mathcal{S} \cup \{P_6\})$*, and let* $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ *be the* $\mathcal{P}(T)$*-free sequence, then the families can be reindexed such that* $\mathcal{F}_1$ *is* $\mathcal{G}_5$*, and the rest of the families are the sets of all cliques.*

Theorem 1.1.17 (Structure of typical T-free graphs for $T \in \mathcal{T}^{\text{npl}} \setminus (\mathcal{S} \cup \{P_6\})$)**.** *Let* $T \in \mathcal{T}^{\text{npl}} \setminus (\mathcal{S} \cup \{P_6\})$*, then almost all* T*-free graphs can be partitioned into* wpn(T) *parts such that one of the parts is a complement of a disjoint union of triangles and stars, and the rest are cliques.*

To summarize, we have the following corollary,

Corollary 1.1.18. *For every tree* T*, almost all* T*-free graphs have a* $P(T)$*-free partition.*

In Chapter 2 we review the definition of the witnessing partition number of a graph H, and provide some additional properties of this function. We also give a more detailed overview of the existing results regarding the typical structure of the H-free graphs for some graphs H.

In Chapter 3 we describe how to obtain a weaker result than in Conjecture 1.1.1. We show that for any graph H, almost all H-free graphs G can be partitioned into $\mathrm{wpn}(H)$ parts $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$ and a set Z such that $(G_1 \setminus Z, G_2 \setminus Z, ..., G_{\mathrm{wpn}(H)} \setminus Z)$ is a $P(H)$-free partition and $|Z| = \mathrm{o}(n)$ where $|V(G)| = n$. Then we reprove the result of Reed and Scott [47] which shows that the above statement can be strengthen so $|Z| \leq n^{1-\varepsilon}$ for some $\varepsilon > 0$ which depends only on H. Finally we show a few general theorems about the structure and properties of almost all H-free graphs from any H. Those theorems are used in the proofs of the typical structure of almost all T-free graphs for a tree T. Note that some of the ideas in those general theorems had already appeared in [47].

In Chapter 4 we prove a theorem regarding the value of the witnessing partition number for any bipartite graph H.

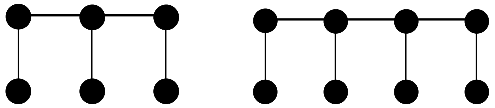
In Chapter 5 reprove the result of Balogh and Butterfield [5] and show that Reed-Scott conjecture 1.1.1 is true for critical graphs. We show that trees without a perfect matching are critical. We prove Theorems 1.1.2, 1.1.4, 1.1.5 and derive Theorems 1.1.3, 1.1.6. Then we prove Reed-Scott conjecture 1.1.1 for any graph H which is a tree with a perfect matching. In Section 5.2.1 we analyze the $P(T)$-free sequence for all trees T and prove Theorems 1.1.8, 1.1.10, 1.1.12, 1.1.14 and 1.1.16. In Section 5.2.2 we prove 1.1.9, 1.1.11, 1.1.13, 1.1.15, and 1.1.17, which imply Corollary 1.1.18.

In Chapter 6 we prove Theorem 1.0.9, and show the Gyárfás-Sumner Conjecture 6 for almost all T-free graphs.

1.2 Additional Notations

For completeness of presentation, we give some additional notations and definitions. We mostly follow the conventions as in [13].

We always consider simple and labeled graphs, that is without parallel edges and loops.



(a) P_3^* (b) P_4^*

Figure 1.4: A sketch of P_3^* and P_4^*.

The set of labels for the vertices is always the set $[|V(G)|] = \{1, 2, ..., |V(G)|\}$. Let $v \in V(G)$, we denote by $\boldsymbol{N_G(v)}$ (respectively, $\boldsymbol{\overline{N}_G(v)}$) the set of vertices which are adjacent (respectively, not adjacent) to v. For $V' \subset V(G)$, we denote by $\boldsymbol{N_G(V')}$= $\cup_{v \in V'} N_G(v)$. Let $\mathbf{deg}_{\boldsymbol{G}}\boldsymbol{(v)}$= $|N_G(v)|$, and let $\overline{\mathbf{deg}}_{\boldsymbol{G}}\boldsymbol{(v)}$= $|\overline{N}_G(v)|$. If there is no confusion regarding the underlaying graph G, we drop the subscript.

Let G be a graph, and let $A, B \subseteq V(G)$ such that $A \cap B = \emptyset$. We denote by $G[A, B]$ the subgraph of G obtained by taking all the edges with one end in A and the other in B.

Let $\boldsymbol{P_k^*}$ be a graph which is a path on k vertices with a leaf adjacent to every vertex on the path. See Figure 1.4 for drawings of P_3^* and P_4^*.

Let G be a graph and let M a matching in G, an **alternating path** in G with respect to M is a path $P = v_1 v_2 v_3 v_4 ... v_k$ which alternates between the edges in the matching and edges not in the matching. More precisely either for each odd $i \in [k]$, $\{v_i, v_{i+1}\} \in M$ and for each even $i \in [k]$, $\{v_i, v_{i+1}\} \notin M$, or vice versa.

A **clique-star** is the complement of a graph which is a disjoint union of a vertex and a complete multi-partite graph with at least two vertices.

Let H be a graph. We denote by $\mathbf{Forb}^{\boldsymbol{w}}\boldsymbol{(H)}$ the family of graphs that do not contain H as a weak subgraph. Similarly we denote by $\mathbf{Forb}^{\boldsymbol{w}}\boldsymbol{(\mathcal{H})}$, the set of graph $\mathcal{H}$, to be the families of all graphs that do not contain H as a weak subgraph for every $H \in \mathcal{H}$.

Let $n, w \in \mathbb{N}$, we use the convention that $\Pi = (\pi_1, \pi_2, ..., \pi_w)$ is the **partition of** $\boldsymbol{[n]}$, that is for all $i \neq j \in [w]$, $\pi_i \cap \pi_j = \emptyset$ and $\cup_{i=1}^{w} \pi_i = [n]$. We say that a partition $\Pi = (\pi_1, \pi_2, ..., \pi_w)$ of $[n]$ is $\boldsymbol{\alpha}$**-almost equal** if for each $i \in [w]$, $|\ |\pi_i| - \frac{n}{w}\ | \leq n^{1-\alpha}$.

We say that a number n_1 is **much smaller** than a number n_2 if $n_1 = \mathrm{o}(n_2)$. Similarly, a number n_1 is **much larger** than a number n_2 if $n_2 = \mathrm{o}(n_1)$.

1.3 Some Useful Theorems

We mention a few results which we use more than once in our proofs.

In many of our proofs we need to bound the number of graphs with some specific properties. Sometimes we use the following bound on binomial coefficients.

Lemma 1.3.1 ([5])**.** *Let $\alpha \in (0, 1/2]$ and $q(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$, then*

$$\sum_{i \leq \alpha n} \binom{n}{i} \leq 2^{q(\alpha)n} \leq 2^{2\alpha \log \frac{1}{\alpha} n}.$$

Here and in all our computations, $\log x$ is $\log_2 x$.

In some cases, we consider families of graphs where each graph is a join (or disjoint union) of other graphs. We can estimate the number of such graphs on $[n]$ vertices using the n-th Bell number which we denote as $\mathrm{Bell}(n)$. The ***n*-th Bell number** is the number of ways to partition the set $[n]$ into any number of parts. We use the following lower and upper bounds on $\mathrm{Bell}(n)$,

Theorem 1.3.2 ([10],[14])**.** *Let $n \in \mathbb{N}$,*

$$\left(\frac{n}{e \ln n}\right)^n < \mathrm{Bell}(n) < \left(\frac{0.792n}{\ln(n+1)}\right)^n.$$

Let $w \in \mathbb{N}$ and let H be a graph and let $(H_1, H_2, ..., H_w)$ be a partition of H. Let G be a graph with $|V(G)| = n$ and let $\Pi := (\pi_1, \pi_2, ..., \pi_w)$ be a partition of $[n]$. Now assume that for each $i \in [w]$, $G[\pi_i]$ contains at least q disjoint copies of the graph H_i for some prime $q \in \mathbb{N}$. For each $i \in [w]$, let $\mathcal{H}_i$ be the set of the disjoint copies of H_i in $G[\pi_i]$. The following lemma shows that we can choose at least q^2 sets of the form $\{H'_1, H'_2, ..., H'_w\}$ where $H'_i \in \mathcal{H}_i$, $i \in [w]$, such that no two sets intersect on more than one graph.

Lemma 1.3.3 ([5])**.** *Let q be a prime, then we can find at least q^2 edge-disjoint cliques of size w in the complete w-partite graph where each part is of size q.*

We denote by $R(c, s)$ the smallest number such that every graph with at least $R(c, s)$ vertices contains either a clique on c vertices or a stable set on s vertices. A famous result by Erdős and Szekeres gives an upper bound on this numbers,

Theorem 1.3.4 (Erdős-Szekeres 35', [29])**.** *For all* $c, s \geq 1$*,*

$$R(c,s) \leq \binom{c+s-2}{c-1}.$$

The following theorems are also useful for our counting arguments.

Theorem 1.3.5 (Markov's inequality)**.** *Let* X *be a non-negative random variable and* $a > 0$*, then*

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

Theorem 1.3.6 (Chernoff Bound [16])**.** *Let* $X_1, X_2, .., X_n$ *be independent random variables with* $\mathbb{P}[X_i = 1] = p_i$ *and* $\mathbb{P}[X_i = 0] = 1 - p_i$*. Let* $X = \sum_{i=1}^{n} X_i$ *with the expectation* $\mathbb{E}[X] = \sum_{i=1}^{n} p_i$*. Then,*

$$\begin{aligned} \mathbb{P}[X \leq E[X] - \lambda] &\leq e^{-\frac{\lambda^2}{2\mathbb{E}[X]}} \\ \mathbb{P}[X \geq E[X] + \lambda] &\leq e^{-\frac{\lambda^2}{2(\mathbb{E}[X]+\lambda/3)}}. \end{aligned}$$

Chapter 2

Structure of H-Free Graphs

In this chapter we give a more detailed review of the known results regarding the structure of H-free graphs for all graphs H and more precise results for some specific graphs H.

2.1 Properties of All Graphs H

Recall that the witnessing partition number of a graph H, $\mathrm{wpn}(H)$, is the maximal sum $s+c$ such that H cannot be partitioned into s stable sets and c cliques. The following properties of the value $\mathrm{wpn}(H)$ were observed in [39] and [46].

Observation 2.1.1. *Let H be a graph, then* $\mathrm{wpn}(H)$ *has the following properties,*

- $\mathrm{wpn}(H) \geq \max\{\chi(H)-1, \sigma(H)-1\}$,
- $\mathrm{wpn}(H) \leq \chi(H)+\sigma(H)-1$,
- $|V(H)| \leq (\mathrm{wpn}(H)+1)^2$.

For example, the last property in the above observation is true because of the following. By the definition of $\mathrm{wpn}(H)$, H *can* be partitioned into $\mathrm{wpn}(H)+1$ stable sets $S_1, S_2, ..., S_{\mathrm{wpn}(H)+1}$ and into $\mathrm{wpn}(H)+1$ cliques $C_1, C_2, ..., C_{\mathrm{wpn}(H)+1}$. Let $i \in [\mathrm{wpn}(H)+1]$, because $|C_i \cap S_j| \leq 1$, for each $j \in [\mathrm{wpn}(H)+1]$, we have that $|C_i| \leq \mathrm{wpn}(H)+1$. Hence $|V(H)| \leq \sum_{i=1}^{\mathrm{wpn}(H)+1} |C_i| \leq (\mathrm{wpn}(H)+1)^2$.

As mentioned earlier many classical results regarding families of graphs which do not contain some graph H as a weak subgraph, which are stated as a function of the chromatic number of H, remain true also for families of H-free graphs if the chromatic number is exchanged by $\tau(H) = \mathrm{wpn}(H) + 1$. One such result is Turán's theorem [56] and its extension by Erdős, Stone and Simonovits [27, 28]. Recall that $\mathrm{Forb}^w(H)$ the family of graphs which do not contain H as a weak subgraph. The family $\mathrm{Forb}(H)$ is the family of H-free graphs. Let $\mathbf{ex}^{\boldsymbol{w}}\boldsymbol{(n, H)}$ be the maximal number of edges in some graph in $(\mathrm{Forb}^w(H))_n$.

Theorem 2.1.2 (Turán [56])**.** *Let K_t be a clique on $t \geq 2$ vertices, then*

$$\mathrm{ex}^w(n, K_t) = \Big(1 - \frac{1}{t-1}\Big)\frac{n^2}{2}.$$

As mentioned the above result was extended by Erdős, Stone and Simonovits [27, 28] for arbitrary graphs H such that $\chi(H) \geq 2$. Their result also gives an approximate version of Turán's Theorem.

Theorem 2.1.3 (Erdős and Stone, Simonovits [27, 28])**.** *Let H be any graph with $\chi(H) \geq 2$, then*

$$\mathrm{ex}^w(n, H) = \Big(1 - \frac{1}{\chi(H) - 1} + \mathrm{o}(1)\Big)\frac{n^2}{2}.$$

In the case of $\mathrm{Forb}(H)$ for some H, a different definition for the extremal graph in $(\mathrm{Forb}(H))_n$, $n \in \mathbb{N}$, is required. Let $\mathbf{ex}\boldsymbol{(n, H)}$ be the maximal number of edges that a graph $G \in (\mathrm{Forb}(H))_n$ may have where there exists a graph $G_0 = (V, E_0)$ with $E \cap E_0 = \emptyset$ and such that $(V, E_0 \cup X)$ does not contain an induced subgraph H for all $X \subseteq E$. This definition was introduced by Prömel and Steger in [46]. The also showed a bound on $\mathrm{ex}(n, H)$.

Theorem 2.1.4 (Prömel and Steger [46])**.** *Let H be a graph. Then*

$$\mathrm{ex}(n, H) = \Big(1 - \frac{1}{\mathrm{wpn}(H)} + \mathrm{o}(1)\Big)\frac{n^2}{2}.$$

Another result we mention concerns the number of graphs in $(\mathrm{Forb}^w(H))_n$ for any graph H and $n \in \mathbb{N}$.

Theorem 2.1.5 (Erdős, Frankl and Rödl [24]). *Let H be a graph and $n \in \mathbb{N}$, then*

$$|(\mathrm{Forb}^w(H))_n| = 2^{\left(1-\frac{1}{\chi(H)-1}+\mathrm{o}(1)\right)\frac{n^2}{2}}.$$

The counterpart theorem for H-free graphs was shown by Prömel and Steger [44].

Theorem 2.1.6 (Prömel and Steger [44]). *Let H be a graph and $n \in \mathbb{N}$, then*

$$|(\mathrm{Forb}(H))_n| = 2^{\left(1-\frac{1}{\mathrm{wpn}(H)}+\mathrm{o}(1)\right)\frac{n^2}{2}}.$$

Note that it is not hard to prove the lower bound in the above theorem (as also in Theorem 2.1.5). Let H be a graph, and let s, c be such that $s + c = \mathrm{wpn}(H)$ and H cannot be partitioned into s stable sets and c cliques. Let $n \in \mathbb{N}$ and $\Pi = \left(\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)}\right)$ be a partition of $[n]$ into $\mathrm{wpn}(H)$ parts of as equal as possible size. Consider a graph G such that $G[\pi_1]$ to $G[\pi_s]$ are stable sets, $G[\pi_{s+1}]$ to $G[\pi_{\mathrm{wpn}(H)}]$ are cliques (the choice of edges between the parts is not restricted). Then by the definition of $\mathrm{wpn}(H)$ and the choice of s and c such a graph is H-free. The number of ways to choose the edges between the parts is $\left(1 - \frac{1}{\mathrm{wpn}(H)} + \mathrm{o}(1)\right)\binom{n^2}{2}$, therefore we get the required lower bound.

For a general hereditary property $\mathcal{F}$ of graphs Alekseev [1] and Bollobás and Thomason [12] showed a similar bound. In order to present the bound we need to define the **colouring number** $\chi_c(\mathcal{F})$ of a property. For each $r \in \mathbb{N}$ and $v \in \{0,1\}^r$, let $\boldsymbol{\mathcal{H}(r, v)}$ be the set of all graphs G such that $V(G)$ can be partitioned into r sets $(\pi_1, \pi_2, ..., \pi_r)$ where for each $i \in [r]$, if $v_i = 0$, then $G[\pi_i]$ is a stable set, and if $v_i = 1$, then $G[\pi_i]$ is clique. Let $\mathcal{F}$ be a hereditary property and let $\boldsymbol{\chi_c(\mathcal{F})}$, be the maximal $r \in \mathbb{N}$ such that $\mathcal{H}(r, v) \subset \mathcal{F}$ for some vector $v \in \{0,1\}^r$.

Note that $\chi_c(\mathrm{Forb}(H)) = \mathrm{wpn}(H)$. Indeed, $\chi_c(\mathrm{Forb}(H)) \geq \mathrm{wpn}(H)$ because by the definition of $\mathrm{wpn}(H)$ there are some s, c such that $s + c = \mathrm{wpn}(H)$ and H cannot be partitioned into s stable sets and c cliques, so any graph G which can be partitioned into s stable sets and c cliques does not contain H and therefore in $\mathrm{Forb}(H)$. On the other hand, $\chi_c(\mathrm{Forb}(H)) \leq \mathrm{wpn}(H)$ because H can be partitioned into s stable sets and c cliques for any s, c such that $s + c = \mathrm{wpn}(H) + 1$, then for any s, c such that $s + c = \mathrm{wpn}(H) + 1$ there is a graph G which can be partitioned into s stable sets and c cliques but is not in $\mathrm{Forb}(H)$.

Theorem 2.1.7 (Alekseev [1], Bollobás and Thomason [12]). *Let $\mathcal{P}$ be a hereditary property of graphs, and suppose $\chi_c(G) = r$. Then*

$$|\mathcal{P}_n| = 2^{\left(1-\frac{1}{r}+o(1)\right)\binom{n^2}{2}}.$$

2.2 Properties of Specific Graphs H

In the following we focus on the structure of the graphs in $\text{Forb}(H)$ for some specific graphs H. We treat separately graphs H with $\text{wpn}(H) = 1$. Note that from Observation 2.1.1, any graph H with $\text{wpn}(H) = 1$ also has $|V(H)| \leq 4$, $\chi(H) \leq 2$ and $\sigma(H) \leq 2$. Hence there are exactly 5 graphs H with $\text{wpn}(H) = 1$, and they are $K_2, S_2, P_3, \overline{P_3}, P_4$. For K_2 (respectively, S_2), there exists only one graph on n vertices that does not contain K_2 (respectively, S_2) and it is the graph with no edges (respectively, a clique). Graphs G that do not contain P_3 (respectively, $\overline{P_3}$) are disjoint unions of cliques (respectively, complete multi-partite graphs), so their number on n vertices is as the number of ways to partition $[n]$. Therefore using the bound on the Bell numbers in Theorem 1.3.2, we have the following.

Theorem 2.2.1. *Let $n \in \mathbb{N}$, $|(\text{Forb}(P_3))_n| = |(\text{Forb}(\overline{P_3}))_n| \leq 2^{n \log n}$.*

Finally we consider the path on 4 vertices P_4.

Theorem 2.2.2 (Seinsche, [53]). *Let G be a P_4-free graph, then either G or $\overline{G}$ is disconnected.*

Let $G \in (\text{Forb}(P_4))_n$, we can encode G by a rooted tree $T(G)$ which has exactly n leaves and no vertices with exactly one child. It is easy to show by induction that in such a tree the number of vertices which are not leaves is at most the number of vertices which are leaves, therefore $T(G)$ has at most $2n$ vertices. We define $T(G)$ recursively as following. We assign the root vertex for G. If G is just one vertex than we are done. Otherwise we use Theorem 2.2.2 about the structure of P_4-free graphs. Assume that G is disconnected, and let $G_1, G_2, ..., G_k$ be its connected components, then we assign k vertices to the components $G_1, G_2, ..., G_k$ and make them the children of the vertex which was assigned to G. Moreover, we colour the edges from a G vertex to $G_1, G_2, ..., G_k$ vertices in red. If G is connected, then we consider the components in $\overline{G}$, assign vertices to them similarly, and colour the edges from a G vertex to the vertices which represents the components in blue. Two different P_4-free graphs are encoded by two different trees.

Theorem 2.2.3 (Cayley's formula [15])**.** *Let $n \in \mathbb{N}$, the number of labeled trees on n vertices is n^{n-2}.*

Using the above theorem, the number of rooted trees on $2n$ vertices with every edge coloured either red or blue is at most $(2n)^{2n-2} \cdot n \cdot 2^{2n-1}$. Therefore we can derive the following theorem.

Theorem 2.2.4. *Let $n \in \mathbb{N}$,*

$$|(\mathrm{Forb}(P_4))_n| \leq 2^{3n \log n}.$$

We mention one more property of P_4-free graphs. This property is a corollary to Theorem 2.2.2 and it will be useful to us in Chapter 6.

Corollary 2.2.5. *Let $G \in \mathrm{Forb}(P_4)$, then G is perfect.*

As described earlier, for a given graph H, every graph G which can be partitioned into s stable sets and c cliques such that $s + c = \mathrm{wpn}(H)$ and H does not have such a partition, is H-free. As mentioned before, this is not the only way in which we can obtain an H-free graph. We recall the Reed-Scott Conjecture.

Conjecture (Reed-Scott [47])**.** *For every graph H, almost all H-free graphs G have a $P(H)$-free partition.*

There are several graphs for which the above conjecture is verified. Firstly, it is known for graphs H which are cliques. Note that for any $t \in \mathbb{N}$, $\mathrm{wpn}(K_t) = t - 1$, this is true because K_t cannot be partitioned into $t - 1$ stable sets, but it can be partitioned into any s stable sets and c cliques such that $s + c = t$. Moreover, K_t can be partitioned into an edge and $t - 2$ vertices, therefore each of the graphs in the $P(H)$-free partition is stable.

Theorem 2.2.6 (Kolaitis, Prömel and Rothschild [38]). *Almost all graphs in* $\mathrm{Forb}(K_t)$ *are* $(t-1)$*-partite graphs.*

Let $\boldsymbol{(\mathcal{F})_{n,m}}$ be the set of all graphs in $\mathcal{F}$ with exactly n vertices and m edges. Osthus, Prömel and Taraz [41] extended the above result in the case of K_3 for different values for the number of edges.

Theorem 2.2.7 ([41]). *Let* $\varepsilon > 0$ *and and let* $m_2(n) = \frac{\sqrt{3}}{4} n^{3/2} \sqrt{\log n}$. *If* $m = \mathrm{o}(n)$ *or* $m \geq (1+\varepsilon) m_2(n)$, *then almost all graphs in* $(\mathrm{Forb}(K_3))_{n,m}$ *are bipartite. If* $n/2 \leq m \leq (1-\varepsilon) m_2(n)$, *then almost all graphs in* $(\mathrm{Forb}(K_3))_{n,m}$ *are not bipartite.*

Balogh, Morris, Samotij and Warnke [9] generalized the above result for any clique.

Theorem 2.2.8 ([9]). *Let* $r \geq 2$, *and let* $\Theta_r = \frac{r-1}{2r} \left(r \cdot \left(\frac{2r+2}{r+2} \right)^{1/r-1} \right)^{2/r+2}$ *and let* $m_r(n) = \Theta_r n^{2-2/r+2} (\log n)^{\frac{1}{\binom{r+1}{2}+1}}$.

For every $r \geq 3$, *there exists a* $d_r(n) = \Theta(n)$, *such that for every* $\varepsilon > 0$, *the following holds. If* $m \leq (1-\varepsilon) d_r(n)$ *or* $m \geq (1+\varepsilon) m_r(n)$, *then almost all graphs in* $(\mathrm{Forb}(K_{r+1}))_{n,m}$ *are* r*-partite. If* $(1+\varepsilon) d_r(n) \leq m \leq (1-\varepsilon) m_r(n)$, *then almost all graphs in* $(\mathrm{Forb}(K_{r+1}))_{n,m}$ *are not* r*-partite.*

Balogh and Butterfield [5] defined a set of **critical graphs**. Let H be a graph and $s, c \in \mathbb{N}$, let $\mathcal{F}(H, s, c)$ denote the set of minimal (by induced containment) graphs F such that H can be covered by s stable sets, c cliques, and F. Balogh and Butterfield called a graph critical if for all s, c such that $s + c = \mathrm{wpn}(H) - 1$ and large enough $n \in \mathbb{N}$, $|\left(\mathrm{Forb}(\mathcal{F}(H, s, c))\right)_n| \leq 2$. As presented later, this means that $\left(\mathrm{Forb}(\mathcal{F}(H, s, c))\right)_n \subseteq \{S_n, K_n\}$. For example, the graph C_4 is critical, but every other cycle of even length is not. Moreover, the graph C_5 is not critical, but every cycle of odd length at least 7 is critical.

Let $\mathcal{W}(H)$ be the collection of all pairs (s, c) such that H cannot partitioned into s stable sets and c cliques where s, c are such that $s + c = \mathrm{wpn}(H)$. Let $\mathcal{Q}(H, s, c)$ be the set of all graphs that can be partitioned into s stable sets and c cliques. Let $\mathcal{Q}(H) = \cup_{(s,c) \in \mathcal{W}(H)} \mathcal{Q}(H, s, c)$.

Theorem 2.2.9 (Balogh and Butterfield [5]). *Let H be a graph with* $\mathrm{wpn}(H) \geq 2$. *Almost all H-free graphs are in $\mathcal{Q}(H)$ if and only if H is critical.*

The easier direction in the above theorem is the necessity. To show it we need the following lemma. The proof is as in [7].

Lemma 2.2.10 (Balogh et al [7], Scheinerman-Zito [49]). *Let $\mathcal{F}$ be a family of graphs and let $n \in \mathbb{N}$. If $|(\mathcal{F})_n| \geq 3$, then $|(\mathcal{F})_n| \geq n-1$.*

Proof. Let $\mathcal{F}$ be a family of graphs and let $n \in \mathbb{N}$. If $|(\mathcal{F})_n| \geq 3$, then there is a graph $G \in (\mathcal{F})_n$ which is not a clique or a stable set. Such a graph G must contain a vertex v, such that $|N(v)|, |\overline{N}(v)| \geq 1$. For every different choice of labels for the vertices in $N(v)$, we get a different graph in $(\mathcal{F})_n$. Therefore $|(\mathcal{F})_n| \geq \binom{n-1}{|N(v)|} \geq n-1$. □

Sketch of the necessity direction in Theorem 2.2.9, [5]. Let H be a non-critical graph. Let $n \in \mathbb{N}$ be large enough, and let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be a partition of $[n]$ into $\mathrm{wpn}(H)$ parts. Let $m(\Pi) = \prod_{1 \leq i < j \leq \mathrm{wpn}(H)} |\pi_i| \cdot |\pi_j|$, that is the number of possible edges between the parts of the partition.

We give an upper bound on the number of possible graphs in $\mathcal{Q}(H)$ that we can get from this partition. Let s, c be such $s + c = \mathrm{wpn}(H)$ and H cannot be partitioned into s stable sets and c cliques. The number of graphs G such that $G[\pi_1]$ to $G[\pi_s]$ are stable sets, $G[\pi_{s+1}]$ to $G[\pi_{\mathrm{wpn}(H)}]$ are cliques is $2^{m(\Pi)}$.

Now, we give a lower bound on the number of possible H-free graphs we get from this partition. Due to the fact that H is not critical, there are c', s' such that $c' + s' = \mathrm{wpn}(H) - 1$ and $|(\mathrm{Forb}(\mathcal{F}(H, c', s')))_k| \geq 3$ for some k large enough. Then by Lemma 2.2.10 it must be the case that $|(\mathrm{Forb}(\mathcal{F}(H, c', s')))_k| \geq k$. Assume without loss of generality that π_1 is the largest part in the partition. The number of graphs G such that $G[\pi_1] \in \mathrm{Forb}(\mathcal{F}(H, c', s'))$, $G[\pi_2]$ to $G[\pi_{s'}]$ are stable sets and $G[\pi_{s'+1}]$ to $G[\pi_{\mathrm{wpn}(H)}]$ are cliques is at least $\frac{n}{\mathrm{wpn}(H)} \cdot 2^{m(\Pi)}$. Note that it is possible that by counting the graphs with respect to a partition we counted much more graphs than there are H-free graphs. As discussed in Subsection 3.3.1, this is actually not the case, and most of the H-free graphs have only one partition.

Therefore the number of H-free graphs is much larger than the number of graphs in $\mathcal{Q}(H)$. $\square$

We mention some additional families of graphs for which Conjecture 1.1.1 is true.

Theorem 2.2.11 (Prömel and Steger [44]). *Almost all graphs in* $\mathrm{Forb}(C_5)$ *are graphs* G *such that either* G *or* $\overline{G}$ *can be partitioned into two sets* V_1, V_2*, such that the first set induces a clique and the second a disjoint union of cliques.*

The odd cycle C_{2k+1} for $k \geq 3$ is critical. Therefore, as a corollary to the result of Balogh and Butterfield 2.2.9, we have the following.

Corollary 2.2.12 (to Theorem 2.2.9). *Almost all graphs in* $\mathrm{Forb}(C_7)$ *can be partitioned either into* 3 *cliques or a stable set and* 2 *cliques. Almost all graphs in* $\mathrm{Forb}(C_{2k+1})$ *for* $k \geq 4$*, can be partitioned into* k *cliques.*

Conjecture 1.1.1 is also true for even cycles.

Theorem 2.2.13 (Prömel and Steger [5, 43]). *Almost all graphs in* $\mathrm{Forb}(C_4)$ *are graphs that can partitioned into a stable set and a clique.*

Theorem 2.2.14 (Reed and Scott [47]). *Almost all graphs in* $\mathrm{Forb}(C_6)$ *can be partitioned into a stable set and a complement of a graph of girth* 5.

Theorem 2.2.15 (Reed and Scott [47]). *Almost all graphs in* $\mathrm{Forb}(C_8)$ *can be partitioned into* 2 *cliques and a graph whose complement is the disjoint union of graphs each of which is a join of a clique and a stable set.*

Theorem 2.2.16 (Reed and Scott [47]). *Almost all graphs in* $\mathrm{Forb}(C_{10})$ *can be partitioned into* 3 *cliques and a graph which is the complement of the disjoint union of stars and cliques.*

Theorem 2.2.17 (Reed and Scott [47]). *Let* $\ell \geq 6$*, almost all graphs in* $\mathrm{Forb}(C_{2\ell})$ *can be partitioned into* $\ell - 2$ *cliques and the complement of a graph which is the disjoint union of stars and cliques of size 3.*

Note that in Theorems 2.2.15-2.2.17 the graphs which are not cliques are taken from the families $\mathcal{G}_6, \mathcal{G}_4, \mathcal{G}_5$, respectively.

Kim, Kühn, Osthus and Townsend [34], using a different approach showed the same result for cycles $C_{2\ell}$ where $\ell \geq 6$, and approximate results for cycles $C_{2\ell}$ where $\ell = 4$ and $\ell = 5$.

Chapter 3

Obtaining Near Partitions

In this chapter we first show that for any graph H, almost all H-free graphs G can be partitioned into wpn(H) parts $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$ such that there is a set Z such that $(G_1 \setminus Z, G_2 \setminus Z, ..., G_{\mathrm{wpn}(H)} \setminus Z)$ is a $P(H)$-free partition and $|Z| = \mathrm{o}(n)$ where $|V(G)| = n$. Then we show that actually the above statement can be strengthened so that $|Z| \leq n^{1-\varepsilon}$ for some $\varepsilon > 0$ which depends only on H. Finally we show a few general theorems which can be applied in the proof of a typical structure of an H-free graph for nearly all graphs H. These theorems are used in the proofs of the typical structure of almost all T-free graphs for a tree T.

3.1 Obtaining Weak Near Partitions

The first step in proving the above mentioned weaker structural result is the celebrated Szemerédi's regularity lemma [55]. Next we present this lemma together with the necessary definitions.

Let G be a graph and let $A, B \subset V(G)$ such that $A \cap B = \emptyset$, then let $\boldsymbol{e(A, B)}$ be the number of edges $\{v_1, v_2\} \in E(G)$ such that $v_1 \in A, v_2 \in B$. The **density** of (A, B) is $d(A, B) = \frac{e(A,B)}{|A| \cdot |B|}$. Let $\varepsilon > 0$, for $A, B \subseteq V(G)$, we say that the pair (A, B) is **$\boldsymbol{\varepsilon}$-regular** if for every two subsets of $A' \subseteq A, B' \subseteq B$, such that $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$, we have that $|d(A, B) - d(A', B')| < \varepsilon$.

That is a regular pair is a pair of sets where every reasonably sized subgraph has a density which is not far from the density of the pair itself.

The regularity lemma says that every dense graph can be partitioned into a constant number of regular pairs and a few leftover edges. Let G be a graph, a vertex partition $V_0, V_1, V_2, ..., V_k$ of $V(G)$ is **ε-regular** for some $\varepsilon > 0$ if (i) $|V_0| \leq \varepsilon n$, (ii) $|V_i| = |V_j|$ for all $1 \leq i, j \leq k$, and (iii) all but εk^2 pairs (V_i, V_j), $1 \leq i < j \leq k$, are ε-regular.

Theorem 3.1.1 (Regularity lemma, Szemerédi 76', [55])**.** *For every $\varepsilon > 0$ and $t \in \mathbb{N}$, there exist $n_0 \in \mathbb{N}$ and T such that for every $n \geq n_0$, every n-vertex graph G admits an ε-regular partition $V_0, V_1, V_2, ..., V_k$ satisfying $t \leq k \leq T$.*

Let $\varepsilon > 0$ and $t \in \mathbb{N}$, and let G be a graph such $|V(G)| = n$ for n large enough. Let $S = S(\varepsilon, t) = (V_0, V_1, V_2, ..., V_k)$ be the ε-regular partition of G that we get from Theorem 3.1.1 for the chosen t. Let $\delta > 0$, the **reduced graph of $\boldsymbol{G}$** with respect to δ and the partition S is the graph with a vertex set $\{v_1, ..., v_k\}$ where each vertex v_i, $i \in [k]$, corresponds to the set V_i in the partition S of G Vertices v_i and v_j, $i \neq j$, are adjacent if and only if (V_i, V_j) is an ε-regular pair and $\delta < d(V_i, V_j) < 1 - \delta$. We denote this reduced graph by $\boldsymbol{R := R(G, S, \delta)}$. Alon et al. [2] showed the following induced version of the famous embedding lemma [37].

Theorem 3.1.2 (Lemma 9, [2])**.** *Let $\delta > 0$ and $h, w \in \mathbb{N}$ then there are $\varepsilon > 0, t \in \mathbb{N}$ and $n_0 = n_0(\delta, h, w, \varepsilon, t) \in \mathbb{N}$ such that the following holds. Let G be a graph with $|V(G)| \geq n_0$ and ε-regular partition $S = S(\varepsilon, t)$. If $R(G, S, \delta)$ contains a K_{w+1}, then for some $s, c \in \mathbb{N}$ such that $s + c = w + 1$, G contains as an induced subgraph any graph H with $|V(H)| \leq h$ and which can be partitioned into c cliques and s stable sets.*

We recall also the famous stability theorem [22, 23, 52]. Let $k, n \in \mathbb{N}$, we denote by $\boldsymbol{T_n(k)}$ the complete multi-partite graph with n vertices and k parts as equal as possible.

Theorem 3.1.3 (Stability Theorem [22, 23, 52])**.** *Let H be a graph. For every $\alpha > 0$ there exists $\beta > 0$ and $n_0 = n_0(\alpha) \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Let G be a graph $G \in (\mathrm{Forb}^w(H))_n$ with $|E(G)| \geq \left(1 - \frac{1}{\chi(H)-1}\right)\frac{n^2}{2} - \beta n^2$, then G can be obtained from $T_n(\chi(H) - 1)$ by changing at most αn^2 edges.*

Now we ready to prove the weaker statement regarding the structure of almost all H-free graphs.

Theorem 3.1.4. *Let H be a graph and let $\lambda > 0$, then there is an $n_0 = n_0(\lambda) \in \mathbb{N}$, such that for all $n \geq n_0$, almost all graphs $G \in (\mathrm{Forb}(H))_n$ have a partition $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ of $V(G) = [n]$ such that there is a set $Z \subset V(G)$ where the following is true. The partition $(G[\pi_1 \setminus Z], G[\pi_2 \setminus Z], ..., G[\pi_{\mathrm{wpn}(H)} \setminus Z])$ is a $P(H)$-free partition and $|Z| \leq \mathrm{wpn}(H) \cdot 2^h \cdot h \cdot \lambda n$, where $h = |V(H)|$.*

Proof. Let $h := |V(H)|$ and $w := \mathrm{wpn}(H)$. For a partition $P = (H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ of H, we define $m(P) := \prod_{1 \leq i < j \leq \mathrm{wpn}(H)} |V(H_i)| \cdot |V(H_j)|$. Let $c(H) := \min_{P \in P(H)} \frac{m(P)-1}{m(P)}$ and let $\mu = \mu(H, \lambda) = -\log c(H) \cdot \left(\frac{\lambda}{2}\right)^2$.

Let $0 < \alpha < \frac{\mu}{16}$. Let $\beta > 0$ and $n_0' \in \mathbb{N}$ be the β and n_0 we get from the stability theorem 3.1.3 for the given α. Let $0 < \delta < \min\left\{\frac{\beta}{8}, \frac{\mu}{16}\right\}$. Let $\varepsilon' > 0$ and n_0'' be the ε and n_0 that we get from Theorem 3.1.2 for the above δ, k, h. Let $0 < \varepsilon < \min\{\frac{\beta}{16}, \frac{\mu}{16}, \varepsilon'\}$ and let $t \in \mathbb{N}$ be such that $\frac{1}{t} \leq \min\left\{\frac{\beta}{8}, \frac{\mu}{16}\right\}$. Let n_0''' and T be the n_0 and T we get from Theorem 3.1.1 for above ε and t. Let $n_0 := \max\{T \cdot n_0', n_0'', n_0'''\}$. Let $\Pi([n])$ be the set of all partitions of $[n]$.

Let $G \in (\mathrm{Forb}(H))_n$ for $n \geq n_0$. By Szemerédi's regularity lemma applied with ε and t as above, the graph G has an ε-regular partition $S = S(\varepsilon, t) = (V_0, V_1, V_2, ..., V_k)$ satisfying $t \leq k \leq T$. Let $\delta > 0$ as above and let $R = R(G, S, \delta)$ be the reduced graph of G. By Theorem 3.1.2, the graph R is $K_{\mathrm{wpn}(H)+1}$-free, otherwise by the definition of $\mathrm{wpn}(H)$, G contains an induced copy of H which contradicts the choice of G. By Turán's Theorem 2.1.2, $|E(R)| \leq \left(1 - \frac{1}{\mathrm{wpn}(H)}\right)\binom{k}{2}$.

Let $(\mathcal{B})_n = (\mathcal{B}(\varepsilon, k, \beta))_n \subset (\mathrm{Forb}(H))_n$ be the set of graphs $G \in (\mathrm{Forb}(H))_n$ which have an ε-regular partition $S = (V_0, V_1, V_2, ..., V_k)$, and such that the reduced graph R contains at most $\left(1 - \frac{1}{\mathrm{wpn}(H)}\right)\binom{k}{2} - \beta k^2$ edges. We bound the number of graphs in $(\mathcal{B})_n$ by counting all the possible partitions of $[n]$ with the above properties. The number of ways to partition $[n]$ into k parts is at most n^k. By our assumptions $|V_0| \leq \varepsilon n$ and $|V_i| \leq \frac{n}{k}$ for each $i \in [k]$. Therefore the number of possible edges of which both ends are contained in some part of the partition is at most $\binom{\varepsilon n}{2} + k\binom{\frac{n}{k}}{2} \leq (\varepsilon + \frac{1}{k})\binom{n}{2}$. Hence the number of choices for graphs on the sets V_i, $i \in [k]$, is $2^{(\varepsilon+\frac{1}{k})\binom{n}{2}} \leq 2^{(\varepsilon+\frac{1}{t})\binom{n}{2}}$. The number of choices for edges between pairs of sets which are not ε-regular or with density smaller than δ or greater than $1-\delta$ is at most $2^{(\varepsilon+\delta)\binom{n}{2}}$. Therefore, we can bound the number of graphs in $(\mathcal{B})_n$ by

$$2^{\left(\left(1-\frac{1}{\mathrm{wpn}(H)}\right)-\beta+\mathrm{o}(1)+\frac{1}{t}+2\varepsilon+\delta\right)\binom{n}{2}}.$$

Note that by our choice of t, ε, δ, we have that $\frac{1}{t}+2\varepsilon+\delta \leq \frac{\beta}{2}$. Therefore if we compare the upper bound on the number of graphs in $(\mathcal{B})_n$ to the lower bound $|(\mathrm{Forb}(H))_n| \geq 2^{\left(1-\frac{1}{\mathrm{wpn}(H)}\right)\binom{n^2}{2}}$ shown in Theorem 2.1.6, we get that $|(\mathcal{B})_n| = \mathrm{o}((\mathrm{Forb}(H))_n)$.

In the following we focus on graphs in $(\mathrm{Forb}(H))_n \setminus (\mathcal{B})_n$. Let $G \in (\mathrm{Forb}(H))_n \setminus (\mathcal{B})_n$, then by the definition of this family of graphs, the corresponding reduced graph R contains at least $\left(1-\frac{1}{\mathrm{wpn}(H)}\right)\binom{k}{2} - \beta k^2$ edges, therefore by the stability theorem 3.1.3, R can be made into $T_k(\mathrm{wpn}(H))$ be changing at most αk^2 edges. Let R' be the resulting graphs after the change of the above αk^2 edges and let $W_1, W_2, ..., W_{\mathrm{wpn}(H)}$ be the partition of $V(R) = V(R')$ into $\mathrm{wpn}(H)$ stable sets. Let $\pi_i = \cup_{v_j \in W_i} V_j$, $i \in [\mathrm{wpn}(H)]$ and let $\Pi(G) = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be the corresponding partition of $[n]$. Note that $G[\pi_i]$, $i \in \mathrm{wpn}(H)$ is not necessarily a stable set, but the number of possible subgraphs of G on each π_i, $i \in [\mathrm{wpn}(H)]$ can be bounded by a function of α, t, ε and δ as follows. By the definitions of R and R', the underlying graph on $G[\pi_i]$ for each $i \in [\mathrm{wpn}(H)]$ is a collection of parts from the ε-regular partition S such that at most αk^2 of the pairs of those parts are ε-regular and have density in $(\delta, 1-\delta)$, let $\mathcal{F}$ be the family of all such possible graphs on each π_i, $i \in [\mathrm{wpn}(H)]$. The number of graphs in $\mathcal{F}$ is at most $2^{\left(\alpha+\frac{1}{t}+\varepsilon+\delta\right)\binom{n^2}{2}} \leq 2^{\frac{\mu}{4}\cdot\binom{n^2}{2}}$, where the inequality is due to the choice of $\alpha, t, \varepsilon, \delta$.

We can conclude that graphs $G \in (\mathrm{Forb}(H))_n \setminus (\mathcal{B})_n$ can be partitioned into $\mathrm{wpn}(H)$ almost equal parts such that the graph induced on each of the parts is a graph from $\mathcal{F}$. Therefore the number of graphs in $(\mathrm{Forb}(H))_n \setminus (\mathcal{B})_n$ is at most

$$2^{\left(1-\frac{1}{\mathrm{wpn}(H)}+\mathrm{o}(1)+\frac{\mu}{4}\right)\binom{n}{2}}.$$

Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be a partition of $[n]$ and let $P = (H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ be a partition of H. Let $(\mathcal{B}(\lambda, \Pi, P))_n \subset \left((\mathrm{Forb}(H))_n \setminus (\mathcal{B})_n\right)$ be the set of graphs

$G \in (\mathrm{Forb}(H))_n \backslash (\mathcal{B})_n$ such that for each $i \in [\mathrm{wpn}(H)]$, $G[\pi_i]$ contains at least λn disjoint copies of H_i, for each $i \in [\mathrm{wpn}(H)]$, let $\mathcal{H}_i$ be the maximum collection of disjoint copies of H_i in $G[\pi_i]$. As before, we want to bound from above the number of graphs in $(\mathcal{B}(\lambda, \Pi, P))_n$. Note that because we consider graphs in $\mathrm{Forb}(H)$, then for each choice of graphs $(H'_1, H'_2, ..., H'_{\mathrm{wpn}(H)})$ such that $H'_i \in \mathcal{H}_i$, $i \in [\mathrm{wpn}(H)]$, there is at least one edge arrangement between those subgraphs that cannot appear otherwise we get a copy of H. Therefore for each choice of graphs $(H'_1, H'_2, ..., H'_{\mathrm{wpn}(H)})$ where $H'_i \in \mathcal{H}_i$, instead of $m(P)$ possible edge arrangements we have at most $m(P) - 1$. Let $K = (H'_1, H'_2, ..., H'_{\mathrm{wpn}(H)})$ be a sequence of graphs where $H'_i \in \mathcal{H}_i$ for each $i \in [\mathrm{wpn}(H)]$ and let $\mathcal{K}$ be a maximum collection of sequences K such that every two different sequences in $\mathcal{K}$ intersect on at most one element. By Lemma 1.3.3, we know that there is a collection $\mathcal{K}$ such that $|\mathcal{K}| \geq \left(\frac{\lambda n}{2}\right)^2$.Therefore the number of graphs in $(\mathcal{B}(\lambda, \Pi, P))_n$ is at most

$$2^{\left(1-\frac{1}{\mathrm{wpn}(H)}+\mathrm{o}(1)+\frac{\mu}{4}\right)\binom{n}{2}} \cdot c(H)^{\left(\frac{\lambda n}{2}\right)^2}$$

$$\leq 2^{\left(1-\frac{1}{\mathrm{wpn}(H)}+\mathrm{o}(1)+\frac{\mu}{4}\right)\binom{n}{2}} \cdot 2^{-\mu \cdot \frac{n^2}{2}}.$$

If we again compare the above bound the lower bound on the number of graphs in $(\mathrm{Forb}(H))_n$ from Theorem 2.1.6, then we can conclude that it is much smaller than the number of graphs in $(\mathrm{Forb}(H))_n$.

Let $G \in \left((\mathrm{Forb}(H))_n \setminus ((\mathcal{B})_n \bigcup (\cup_{\Pi \in \Pi([n]), P \in P(H)} (\mathcal{B}(\lambda, \Pi, P))_n)\right)$, and let $\Pi(G)$ be a partition of $[n]$ as above. We define the set Z_i for each $i \in [\mathrm{wpn}(H)]$, to be the union of the vertex set of all disjoint copies of all induced subgraph H' of H, such that $G[\pi_i]$ does not contain at least λn disjoint copies of H'. By setting $Z = \cup_{i=1}^{\mathrm{wpn}(H)} Z_i$, we get the required partition. Indeed, if $(G[\pi_1 \setminus Z], G[\pi_2 \setminus Z], ..., G[\pi_{\mathrm{wpn}(H)} \setminus Z])$ is not $P(H)$-free partition then there is a partition $(H_1, H_2, ..., H_{\mathrm{wpn}(H)})$ of H such that for each $i \in [\mathrm{wpn}(H)]$, H_i is an induced subgraph in $G[\pi_i \setminus Z]$, but then, by the definition of Z, each $G[\pi_i \setminus Z]$ contains at least λn disjoint copies of H_i, and this is a contradiction to the choice of G. There are at most 2^h subgraphs of H, each of them has size at most h and there are at most λn disjoint copies of each such subgraph, therefore $|Z| \leq \mathrm{wpn}(H) \cdot 2^h \cdot h \cdot \lambda n$, as required. □

3.2 Obtaining Better Near Partitions

We start by presenting a result due to Alon, Balogh, Bollobás and Morris [2]. This result is a crucial initial step in our arguments. Firstly we show how we can use this result to strengthen the partition of almost all H-free graphs for all graphs H. We show that almost all H-free graphs G can be partitioned into $\mathrm{wpn}(H)$ parts $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$ where there is a set Z such that $(G_1 \setminus Z, G_2 \setminus Z, ..., G_{\mathrm{wpn}(H)} \setminus Z)$ is a $P(H)$-free partition and there is an $\varepsilon > 0$ such that $|Z| \leq n^{1-\varepsilon}$.

Before we can present the result of Alon et al [2] we need a few definitions. Let $\boldsymbol{U(k)}$ be a bipartite graph with parts $A \cong [2]^k$ and $B \cong [k]$ and edges between a vertex $a \in A$ and $b \in B$ if and only if $b \in a$. We say that G contains a **copy of $\boldsymbol{U(k)}$** if there are $A, B \subseteq V(G)$ such that $A \cap B = \emptyset$ and $G[A, B]$ is isomorphic to $U(k)$. Alon et. al. [2] showed the following theorem. Note that part (iii) does not appear in the statement of their main theorem, but can be derived from the proof, see [47].

Theorem 3.2.1 (Alon et al [2])**.** *Let $\mathcal{F}$ be a hereditary property of graphs with a colouring number $\chi_c(\mathcal{F}) = r$ and let $\delta > 0$. Then there exist constants $k = k(\mathcal{F}) \in \mathbb{N}$, $\varepsilon = \varepsilon(\mathcal{F}) > 0$ and $b = b(\mathcal{F}, \delta) \in \mathbb{N}$ such that the following holds.*

For almost all graph $G \in \mathcal{F}$, there exists a partition $(\pi_1, \pi_2, ..., \pi_r)$ of $V(G) = [n]$ such that $\pi_i = \pi_i' \dot\cup X_i$ for each $i \in [r]$, and a set $B \subset V(G)$ of at most b vertices such that,

(i) $G[\pi_i']$ does not contain a copy of $U(k)$ for every $i \in [r]$,

(ii) $| \cup_{i=1}^{r} X_i| \leq n^{1-\varepsilon}$,

(iii) for every vertex $v \in \pi_i$, $i \in [r]$, there is a vertex $b \in B$, $|(N(v) \triangle N(b)) \cap \pi_i)| \leq \delta n$.

We denote $X := \cup_{i=1}^{\mathrm{wpn}(H)} X_i$. The following theorem is due to Reed and Scott [47] and can be derived from the results in [2].

Theorem 3.2.2 (Reed-Scott [47])**.** *For every graph H and constant $\xi > 0$, there are $\rho = \rho(H, \xi) > 0$ and $b = b(H, \xi) \in \mathbb{N}$, such that the following holds.*

For almost all H-free graphs G, there exists a partition $(\pi_1, \pi_2, ..., \pi_{\text{wpn}(H)})$ of $V(G) = [n]$ such that $\pi_i = \pi'_i \dot{\cup} Z_i$ for each $i \in [\text{wpn}(H)]$, and a set $B \subset V(G)$ of at most b vertices such that,

(I) the partition $(G[\pi'_1], G[\pi'_2], ..., G[\pi'_{\text{wpn}(H)}])$ is a $P(H)$-free partition,

(II) $|\cup_{i=1}^{\text{wpn}(H)} Z_i| \leq n^{1-\rho}$,

(III) for every $i \in [\text{wpn}(H)]$ and vertex $v \in \pi_i$, there is a vertex $b \in B$ such that

$$\left| \big(N(v) \triangle N(b)\big) \cap \pi_i \right| \leq \xi n,$$

(IV) for every $i \in [\text{wpn}(H)]$, we have

$$\left| |\pi_i| - \frac{n}{\text{wpn}(H)} \right| \leq n^{1-\frac{\rho}{4}}.$$

Reed and Scott proved the above theorem using Theorem 3.2.1 together with the following theorem.

Theorem 3.2.3 ([2])**.** *For each $k \in \mathbb{N}$ there exists $\varepsilon = \varepsilon(k) > 0$ such that there are at most $2^{\ell^{2-\varepsilon}}$ distinct graphs which do not contain a copy of $U(k)$ on the vertex set $[\ell]$.*

For completeness we reprove the Reed-Scott theorem.

Proof of Theorem 3.2.2. Let H be a graph and let $\xi > 0$, let $\delta \in \left(0, \min\{\xi, \frac{1}{2^4 \text{wpn}(T)}\}\right)$, from Theorem 3.2.1 applied with $\text{Forb}(H)$ and δ, we get $k, b \in \mathbb{N}$ and $\varepsilon > 0$ such that almost all graphs in $\text{Forb}(H)$ have a partition $\Pi(G)$ of their vertex set with respect to which they have properties (i)-(iii). We show that almost all graphs $G \in \text{Forb}(H)$ also have properties $(I) - (III)$ with respect to $\Pi(G)$. Note that property (III) is an immediate consequence to property (iii) and the choice of δ. Let $\Pi([n])$ be the set of all partitions of $[n]$.

We start from bounding the number of graphs in $(\text{Forb}(H))_n$. We do that by bounding the number of graphs in $(\text{Forb}(H))_n$ with a partition with respect to which they have properties (i)-(iii) as in Theorem 3.2.1. There are at most $\text{wpn}(H)^n$ ways to partition $[n]$ into $\text{wpn}(H)$ parts. Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\text{wpn}(H)})$ be a partition on $[n]$, and let $n_i := |\pi_i|$, $i \in [\text{wpn}(H)]$. For each $i \in [\mathrm{wpn}(H)]$, we partition further $\pi_i = \pi_i' \cup X_i$. There are at most 2^n ways to do so. There are at most $2^n \cdot 2^{bn}$ ways to choose the vertices in B and to choose their neighbourhoods in the graph.

By our assumptions, $G[\pi_i']$ does not contain a copy of $U(k)$. By Theorem 3.2.3 there is a $\varepsilon' = \varepsilon'(k) > 0$ such that there are at most $2^{\ell^{2-\varepsilon'}}$ graphs which do not contain a copy of $U(k)$ on $[\ell]$. Therefore we have at most $2^{n^{2-\varepsilon'}}$ ways to choose the graphs on π_i'. Let $i \in [\mathrm{wpn}(H)]$ and let $x \in X_i$, then using part (iii) in Theorem 3.2.1, there is a vertex $b \in B$ such that the neighbourhood of x in π_i is similar to the neighbourhood of b in π_i. There are b ways to choose such a vertex in B and there are at most $\binom{n_i}{\delta n}$ ways to choose the neighbourhood of x in π_i which differs from the neighbourhood of b. Therefore the number of ways to choose the neighbourhood of x in π_i is at most $b\binom{n_i}{\delta n}$. If we take the product over all vertices x in X, then using the bound in 1.3.1, the number of ways to choose the neighbourhood of them inside their corresponding parts is at most $b^{|X|} 2^{2\delta \log \frac{1}{\delta} n |X|}$. Hence using the bound on the size of X in part (ii) of Theorem 3.2.1 we deduce that the number of graphs in $(\mathrm{Forb}(H))_n$ is at most

$$2^{n^{2-\varepsilon'} + 2\delta \log \frac{1}{\delta} n^{2-\varepsilon} + \left(1 - \frac{1}{\mathrm{wpn}(H)}\right)\binom{n}{2} + O(n)}.$$

Next, using the above general bound on the number of graphs in $(\mathrm{Forb}(H))_n$, we bound now the number of graphs which do not have properties (I), (II) or (IV). First we consider properties (I) and (II). We proceed as we did in the proof of Theorem 3.1.4.

Let $\alpha \in \left(0, \min\{\frac{\varepsilon'}{16}, \frac{\varepsilon}{16}\}\right)$, let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be a partition of $[n]$ and let $P = (H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ be a partition of H. Let $(\mathcal{B}(\alpha, \Pi, P))_n \subset (\mathrm{Forb}(H))_n$ be the set of graphs $G \in (\mathrm{Forb}(H))_n$ such that for each $i \in [\mathrm{wpn}(H)]$, $G[\pi_i]$ contains at least $n^{1-\alpha}$ disjoint copies of H_i. Let, for each $i \in [\mathrm{wpn}(H)]$, $\mathcal{H}_i$ be the maximum collection of disjoint H_i in $G[\pi_i]$. We want to bound the number of graphs in $(\mathcal{B}(\alpha, \Pi, P))_n$. Note that because $G \in \mathrm{Forb}(H)$, for each choice $(H_1', H_2', ..., H_{\mathrm{wpn}(H)}')$ of graph such that $H_i' \in \mathcal{H}_i$, $i \in [\mathrm{wpn}(H)]$, there is at least one edge arrangement between those sets that cannot appear, otherwise we get a copy of H. Let $K = (H_1', H_2', ..., H_{\mathrm{wpn}(H)}')$ be a sequence of graphs where $H_i' \in \mathcal{H}_i$ for each $i \in [\mathrm{wpn}(H)]$ and let $\mathcal{K}$ be a maximum collection of sequences K such that every two different sequences in $\mathcal{K}$ intersect on at most one element. Using Lemma 1.3.3, $|\mathcal{K}| \geq \frac{n^{2-2\alpha}}{4}$. Therefore the number of graphs in $(\mathcal{B}(\alpha, \Pi, P))_n$ is at most

$$2^{n^{2-\varepsilon'}+2\delta \log \frac{1}{\delta} n^{2-\varepsilon}+\left(1-\frac{1}{\mathrm{wpn}(H)}\right)\binom{n}{2}+\mathrm{O}(n)} \cdot 2^{-c(H)n^{2-2\alpha}}$$

where $c(H) > 0$ is a constant which depends only on the graph H. By the choice of α and Theorem 2.1.6 the above is much smaller than the number of graphs in $(\mathrm{Forb}(H))_n$.

Let $G \in (\mathrm{Forb}(H))_n \setminus (\cup_{\Pi\in\Pi([n]),P\in P(H)}(\mathcal{B}(\alpha, \Pi, P))_n)$, and let $\Pi(G)$ be a partition of $V(G)$ as above. For each $i \in [\mathrm{wpn}(H)]$, let Z_i be the union of all disjoint copies of all induced subgraphs H' of H, such that $G[\pi_i]$ does not contain at least $n^{1-\alpha}$ disjoint copies of H'. Let $Z = \cup_{i=1}^{\mathrm{wpn}(H)} Z_i$. By the choice of G, the partition $(G[\pi_1 \setminus Z], G[\pi_2 \setminus Z], ..., G[\pi_{\mathrm{wpn}(H)} \setminus Z])$ is $P(H)$-free. Moreover, there are 2^h different subgraphs of H, so $|Z| \leq 2^h \cdot n^{1-\alpha}$. Let $\rho = \alpha/2$, then G has properties (I)-(II) with respect to $\Pi(G)$ and Z chosen as above.

To finish the proof, we show that the number of graphs G in $(\mathrm{Forb}(H))_n$ that do not have a partition $\Pi(G) = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ as above with property (IV) is much smaller than the number of graphs in $(\mathrm{Forb}(H))_n$. Let $|\pi_i| = n_i = \frac{n}{\mathrm{wpn}(H)} + a_i$ for $a_i \in \mathbb{Z}$, $i \in [\mathrm{wpn}(H)]$. The number of edges between the parts is at most

$$\binom{n}{2} - \sum_{i=1}^{\mathrm{wpn}(H)} \binom{n_i}{2} = \binom{n}{2} - \sum_{i=1}^{\mathrm{wpn}(H)} \binom{\frac{n}{\mathrm{wpn}(H)} + a_i}{2} \leq \binom{n}{2} - \frac{n^2}{2\,\mathrm{wpn}(H)} - \frac{1}{2}\sum_{i=1}^{\mathrm{wpn}(H)} a_i^2.$$

Let $\rho = \alpha/2$ as before and assume that there is an index $i \in [\mathrm{wpn}(H)]$ such that $a_i \geq n^{1-\frac{\rho}{4}}$, then the number of such graphs is at most

$$2^{n^{2-\varepsilon'}+2\delta \log \frac{1}{\delta} n^{2-\varepsilon}+\left(1-\frac{1}{\mathrm{wpn}(H)}\right)\binom{n}{2}+\mathrm{O}(n)-n^{2-\frac{\rho}{2}}}.$$

By the choice of ρ and Theorem 2.1.6 the above is much smaller than the number of graphs in $(\mathrm{Forb}(H))_n$. □

3.3 General Tools

In this section we present a collection of tools which can be used in proving an exact structural theorem of almost all H-free graphs for any graph H. All the following tools are used in the proof of the exact structure of all almost all T-free graphs for any tree T.

Let H be a graph and let $\xi > 0$ be a constant to be defined later. Let $\rho = \rho(H, \xi)$ and let B be the constant and the set, respectively, that we get from Theorem 3.2.2 applied with H and ξ. Let $Z = \cup_{i=1}^{\mathrm{wpn}(H)} Z_i$. Let $n \in \mathbb{N}$ be large enough and let Π be a $\rho/4$-almost equal partition of $[n]$ into $\mathrm{wpn}(H)$ parts. In the following when we refer to properties (I),(II),(III) and (IV), we always mean those properties from Theorem 3.2.2.

Let $\rho > 0$ and Π be a $\rho/4$-almost equal partition of $[n]$. A graph $G \in \mathrm{Forb}(H)$ is **Π-conformal** if it has properties (I)-(III) with respect to Π. Let $\boldsymbol{\mathcal{RS}} \subseteq \mathrm{Forb}(H)$ be the set of all graphs which are Π-conformal with respect to some partition Π as above.

A graph $G \in (\mathcal{RS})_n$ is **good** if there is some partition Π with respect to which it is Π-conformal and the set Z is empty (in this case, requiring property (II) is redundant). A graph $G \in (\mathcal{RS})_n$ is **Π-good** if it is Π-conformal and has the set Z empty with respect to Π.

A graph $G \in (\mathcal{RS})_n$ is **bad** if for every partition Π with respect to which it is Π-conformal, G is not Π-good. In other words, there is no partition Π such that G has properties (I)-(III) and Z is empty with respect to this partition. A graph $G \in (\mathcal{RS})_n$ is **Π-bad** if it is Π-conformal and bad.

Our goal is to show that the number of bad graphs is much smaller than the number of good graphs. We do it by choosing $n \in \mathbb{N}$ large enough, fixing a ρ-almost equal partition Π of $[n]$ for some $\rho > 0$ and then showing that the number of Π-bad graphs is much smaller than the number of Π-good graphs. Firstly we need to establish that indeed showing that the number of Π-bad graphs is much smaller than the number of Π-good graphs, for every suitable partition Π, implies that the number of bad graphs is much smaller than the number of good graphs. We do that by showing in the next subsection that the number of Π-good graphs with respect to all suitable partitions Π is not much larger than the number of good graphs.

3.3.1 Partitions

In order to present the main theorem of this subsection, we need the following definitions and claims.

Lemma 3.3.1 (Lemma 6 for $r = 2$, [2])**.** *For each $k \in \mathbb{N}$, there exists $K = K(k) \in \mathbb{N}$ such that the following holds. Let G be a graph which contains a copy of $U(K)$, then it contains a copy of $U(k)$ such that $V(U(k)) = A \cup B$, $G[A, B] = U(k)$, and each of $G[A]$ and $G[B]$ is either a clique or a stable set.*

Lemma 3.3.2. *Let H be a graph, then for any $s, c \in \mathbb{N}$ such that $s + c = \mathrm{wpn}(H) - 1$, H can be partitioned into s stable sets, c cliques and either a bipartite graph, or a complement of a bipartite graph or a graph which can be partitioned into a stable set and a clique.*

Proof. Let $s, c \in \mathbb{N}$ such that $s + c = \mathrm{wpn}(H) - 1$, then we argue that it is possible to partition H into s stable sets, c cliques and a bipartite graph. The proof for the case of complement of a bipartite graph or a graph which can be partitioned into a stable set and a clique is similar. By the definition of $\mathrm{wpn}(H)$, H can be partitioned into $s + 2$ stable $(S_1, S_2, ..., S_{s+2})$ sets and c cliques $(C_1, C_2, ..., C_c)$ $(s + 2 + c = \mathrm{wpn}(H) + 1)$. Then s stable sets $(S_1, S_2, ..., S_s)$ together with c cliques $(C_1, C_2, ..., C_c)$ and $H[V(S_1) \cup V(S_2)]$ is the required partition of H. □

Let H be a graph and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence. Note that because the families in a $P(H)$-free sequence are infinite, then using Theorem 1.3.4 we can deduce that each of the families $\mathcal{F}_i$ contains either all the stable sets or all the cliques. Moreover, the graph U where $V(U) = A \cup B$, $U[A, B] = U(h)$ and $U[A]$, $U[B]$ are stable sets contains as an induced subgraph every induced bipartite graph of H. A similar argument is true if one of the $U[A]$ and $U[B]$ or both are cliques. Hence we can use the above lemma to obtain the following property of $P(H)$-free sequences.

Corollary 3.3.3. *Let H be a graph with $|V(H)| = h$ and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence. Then for each $i \in [\mathrm{wpn}(H)]$, $\mathcal{F}_i$ does not contain a graph U such that $V(U) = A \cup B$, each of $U[A]$ and $U[B]$ is either a clique or a stable set and $U[A, B] = U(h)$.*

Let H be a graph and let $\mathscr{G}$ be the collection of all good graphs in $\mathcal{RS} \subseteq \text{Forb}(H)$. Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ be a $P(H)$-free sequence, we say that a Π-good graph $G \in \mathscr{G}$ **extends** the sequence $(G_1, G_2, ..., G_{\text{wpn}(H)})$, $G_i \in \mathcal{F}_i$, $i \in [\text{wpn}(H)]$, if G is obtained by adding edges between the graphs G_i, $i \in [\text{wpn}(H)]$.

Lemma 3.3.4. *Let $k, n \in \mathbb{N}$, $\rho > 0$ and $\Pi = (\pi_1, \pi_2, ..., \pi_{\text{wpn}(H)})$ be a ρ-almost equal partition of $[n]$, let $\varepsilon \in \left(0, \frac{1}{2}\right)$ and $C \geq k$. Let $\mathcal{S} = (G_1, G_2, ..., G_{\text{wpn}(H)})$ be a sequence of graphs such that $V(G_i) = \pi_i$ for each $i \in [\text{wpn}(H)]$. Then the number of graphs G which extend the sequence $\mathcal{S}$ and do not have the following property is $\text{o}(n)$ of the number of all the graphs which extend $\mathcal{S}$. For any two indices $i \neq j$ and for all sets $W_i \subseteq V(G_i)$ and $W_j \subseteq V(G_j)$ such that $|W_i|, |W_j| \geq n^{\frac{1}{2}+\varepsilon}$ or $|W_i| \geq C$ and $|W_i| \geq |V(G_j)| - n^{\frac{1}{2}+\varepsilon}$ there are subsets $A_i \subset W_i$ and $B_j \subset W_j$ such that $G[A_i, B_j] = U(k)$.*

Proof. Let $\mathcal{S}$ be the sequence as above. We obtain a Π-good graph G which extends $\mathcal{S}$ by choosing edges between the parts with probability $\frac{1}{2}$.

Let $W_i \subseteq V(G_i)$ and $W_j \subseteq V(G_j)$, $i \neq j$, be two subsets such that $|W_i|, |W_j| \geq n^{\frac{1}{2}+\varepsilon}$. The probability that for any $A_i \subset W_i$ and $B_j \subset W_j$, $G[A_i, B_j] \neq U(k)$, is at most

$$\begin{aligned} &\binom{n^{\frac{1}{2}+\varepsilon}}{k} \cdot \binom{n^{\frac{1}{2}+\varepsilon}}{2^k} \left(\frac{2^{k \cdot 2^k} - 1}{2^{k \cdot 2^k}}\right)^{\frac{n^{\frac{1}{2}+\varepsilon}}{k} \cdot \frac{n^{\frac{1}{2}+\varepsilon}}{2^k}} \\ &\leq 2^{2(n^{\frac{1}{2}+\varepsilon})} \cdot 2^{-c(k)n^{1+2\varepsilon}}, \end{aligned}$$

where $c(k) > 0$ is a constant which depends only on k. Therefore the expected number of sets W_i, W_j as above is at most

$$\begin{aligned} &\text{wpn}(H)^2 \cdot \binom{n}{n^{\frac{1}{2}+\varepsilon}}^2 \cdot 2^{2(n^{\frac{1}{2}+\varepsilon})} \cdot 2^{-c(k)n^{1+2\varepsilon}} \\ &\leq 2^{6n} \cdot 2^{-c(k)n^{1+2\varepsilon}} = \text{o}(n). \end{aligned}$$

Hence the probability that a graph G which extends $\mathcal{S}$ has that for any $W_i \subseteq V(G_i)$ and $W_j \subseteq V(G_j)$, $i \neq j$ such that $|W_i|, |W_j| \geq n^{\frac{1}{2}+\varepsilon}$, for any $A_i \subset W_i$ and $B_j \subset W_j$, $G[A_i, B_j] \neq U(k)$ is $\text{o}(n)$.

Let $W_i \subseteq V(G_i)$ and $W_j \subseteq V(G_j)$, $i \neq j$, be two subsets such that $|W_i| \geq |V(G_i)| - n^{\frac{1}{2}+\varepsilon}$ and $|W_j| \geq C$. The probability that for any $A_i \subset W_i$ and $B_j \subset W_j$, $G[A_i, B_j] \neq U(k)$, is at most

$$\binom{C}{k}\binom{|V(G_i)| - n^{\frac{1}{2}+\varepsilon}}{2^k} \cdot \left(\frac{2^{k \cdot 2^k} - 1}{2^{k \cdot 2^k}}\right)^{\frac{|V(G_i)|-n^{\frac{1}{2}+\varepsilon}}{k} \cdot \frac{C}{2^k}}$$
$$\leq 2^{2^{2k+2}\log n} \cdot 2^{-c(k,C,H)n},$$

where $c(k, C, H) > 0$ is a constant which depends only on k, C, H. Therefore the expected number of sets W_i, W_j as above is at most

$$\mathrm{wpn}(H)^2 \cdot \binom{n}{n^{\frac{1}{2}+\varepsilon}}^2 \cdot 2^{2^{2k+2}\log n} \cdot 2^{-c(k,C,H)n}$$
$$\leq 2^{4n^{\frac{1}{2}+\varepsilon}\log n} \cdot 2^{-c(k,C,H)n} = \mathrm{o}(n).$$

Hence the probability for the existence of such sets is $\mathrm{o}(n)$. This completes the proof. $\square$

Let F be a graph, we say that a graph $F \cup \{v\}$ is a result of randomly adding a vertex v to F if it is an outcome of adding every edge between v and $V(F)$ with probability $\frac{1}{2}$.

A hereditary family of graphs $\mathcal{F}$ is **stable** if for all $t \in \mathbb{N}$ and almost all graphs $F \in \mathcal{F}$, if we randomly add a vertex v then almost surely there are subgraphs $F_1, F_2, ..., F_t$ of $F \cup \{v\}$ such that for any $i \neq j \in [t]$, $V(F_i) \cap V(F_j) = \{v\}$ and $F_i \notin \mathcal{F}$ for all $i \in [t]$. For example, if $\mathcal{F}$ is the family of all cliques, then if we randomly add a vertex v to a large enough clique $F \in \mathcal{F}$, then almost surely the non-neighbourhood of v is large enough so we get the required subgraphs $F_1, F_2, ..., F_t$.

Let $w \in \mathbb{N}$, a sequence of families $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_w)$ is **properly arranged** if for any $i \neq j \in [w]$, either $\mathcal{F}_i = \mathcal{F}_j$ or almost all graphs $G_i \in \mathcal{F}_i$ cannot made into a graph $G_j \in \mathcal{F}_j$ by deleting a constant number of vertices. For example, a sequence of families $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_w)$ which is *not* properly arranged is as follows. The family $\mathcal{F}_1$ is the family of all graphs which are cliques or a disjoint union of a clique and a singleton vertex. The rest of the families $\mathcal{F}_i$, $i \geq 2$, are the families of all cliques.

Theorem 3.3.5. *Let H be a graph and assume that all $P(H)$-free sequences $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ are properly arranged and moreover for each $i \in [\text{wpn}(H)]$, $\mathcal{F}_i$ is stable. Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ be a $P(H)$-free sequence, $n \in \mathbb{N}$ be large enough and $\rho > 0$, then almost all good graphs G have the following property. If G is Π-good and Π'-good so $G[\pi_i], G[\pi'_i] \in \mathcal{F}_i$, then there is permutation σ so $\pi_i = \pi'_{\sigma(i)}$ for each $i \in [\text{wpn}(H)]$.*

Proof. Let G be a good graph and let $\Pi = (\pi_1, \pi_2, ..., \pi_{\text{wpn}(H)})$, $\Pi' = (\pi'_1, \pi'_2, ..., \pi'_{\text{wpn}(H)})$ be ρ-almost equal partitions of $[n]$ so G is Π-good and Π'-good. Moreover, let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ be the $P(H)$-free sequence so $G[\pi_i], G[\pi'_i] \in \mathcal{F}_i$ for each $i \in [\text{wpn}(H)]$. Assume to the contrary that $\Pi' \neq \Pi$.

Let $h = |V(H)|$, let $k \geq h$, and let $K = K(k)$ be the constant from Lemma 3.3.1. From Corollary 3.3.3, we know that for each $i \in [\text{wpn}(H)]$, $G[\pi_i]$ and $G[\pi'_i]$ do not contain a copy of $U(K)$.

Let $i \in [\text{wpn}(H)]$ be such that $|\pi_1 \cap \pi'_i| \geq n^{\frac{1}{2}+\varepsilon}$ for some $\varepsilon > 0$. Such an index exists because both Π and Π' are ρ-almost equal partitions. Then it must be the case that $|\pi_1 \triangle \pi'_i| \leq 2n^{\frac{1}{2}+\varepsilon}$, otherwise, by Lemma 3.3.4, $G[\pi_1]$ or $G[\pi'_i]$ contains $U(K)$. But if this is the case then actually it must be that $|\pi_1 \triangle \pi'_i| < 2 \cdot 2^k$, because otherwise, again by Lemma 3.3.4, $G[\pi_1]$ or $G[\pi'_i]$ contains $U(K)$.

By our assumptions, $G[\pi_1] \in \mathcal{F}_1$, then because every $P(H)$-free sequence is properly arranged it must be the case that $G[\pi'_i] \in \mathcal{F}_1$. Let $v \in \pi_1 \triangle \pi'_i$, assume without loss of generality that $v \in \pi'_i \setminus \pi_1$. By our assumptions the family $\mathcal{F}_1$, $i \in [\text{wpn}(H)]$, is stable, therefore there are subgraphs $F_1, F_2, ..., F_t$ of $G[\pi_1 \cup \{v\}]$ such that for any $i_1 \neq i_2 \in [t]$, $V(F_{i_1}) \cap V(F_{i_2}) = \{v\}$ and $F_j \notin \mathcal{F}_1$ for all $j \in [t]$ and moreover $t \geq 2 \cdot 2^k$. By the choice of t and the fact that $|\pi_1 \triangle \pi'_i| < 2 \cdot 2^k$, there must be a graph F_j, $j \in [t]$, such that $V(F_j) \setminus \{v\} \subset \pi'_i$, but because $F_j \notin \mathcal{F}_1$, then $G[\pi'_i \cup \{v\}] \notin \mathcal{F}_1$ and therefore, $v \notin \pi'_i$, contradicting the choice of v. Hence $\pi'_i \setminus \pi_1 = \emptyset$. By a symmetric argument, $\pi_1 \setminus \pi'_i = \emptyset$. Therefore $\pi_1 = \pi'_i$.

Repeating the same argument as above, we can conclude that there is mapping $f : [\text{wpn}(H)] \to [\text{wpn}(H)]$ such that $\pi_i = \pi'_{f(i)}$. Moreover there is an permutation of the parts in the partition Π' so $f(i) = i$ for each $i \in [\text{wpn}(H)]$. $\square$

3.3.2 Additional preliminaries

Let H be a graph and let $\boldsymbol{\Phi(H)}$ be the set of all $P(H)$-free sequences. Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be a partition of $[n]$. Then we define the following values.

$$\boldsymbol{m := m(\Pi)} = \prod_{1 \le i < j \le \mathrm{wpn}(H)} |\pi_i| \cdot |\pi_j|,$$

$$\boldsymbol{F := F(H, \Pi)} = \max_{(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)}) \in \Phi(H)} \prod_{i=1}^{\mathrm{wpn}(H)} |\,(\mathcal{F}_i)_{|\pi_i|}\,|.$$

Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)}) \in \Phi(H)$, by the definition of a $P(H)$-free sequence, for every choice of graphs $G_i \in (\mathcal{F}_i)_{|\pi_i|}$, $i \in [\mathrm{wpn}(H)]$, and any choice of the edges between the graphs $(G_1, G_2, ..., G_{\mathrm{wpn}(H)})$, the resulting graph is a Π-good graph. Therefore we can obtain the following lower bound. In many of the following arguments when we want to show that almost all H-free graphs do not have some property P, we compare the number of graphs with property P to the following lower bound.

Observation 3.3.6 (Lower bound on Π-good graphs)**.** *Let H be a graph, and let Π be a partition of $[n]$. The number of Π-good graphs in $(\mathrm{Forb}(H))_n$ is at least,*

$$2^m \cdot F.$$

Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be a ρ-almost equal partition of $[n]$ for some $\rho > 0$ and let $\lambda > 0$ and $i \in [\mathrm{wpn}(H)]$. Let $G \in \mathcal{RS}$ be a Π-conformal graph. A subgraph J of H is $\boldsymbol{(\Pi, \lambda, i)}$**-common** in G if there is a set $\mathcal{J}$ of disjoint copies of graphs isomorphic to J in $G[\pi_i]$ and $|\mathcal{J}| \ge n^{1-\lambda}$.

Let Π be a partition as above and let $h \in \mathbb{N}$, $\lambda > 0$ and $\mu > 0$. A graph $G \in \mathcal{RS}$ is $\boldsymbol{(\Pi, h, \lambda, \mu)}$**-exceptional** if G is Π-conformal and there is a graph J decomposable into $J_1, J_2, ..., J_{\mathrm{wpn}(H)}$ such that for each $i \in [\mathrm{wpn}(H)]$, $|V(J_i)| \le h$, J_i is (Π, λ, i)-common in G and moreover the following is true for G. Let $\mathcal{J}_i$, $i \in [\mathrm{wpn}(H)]$, be the maximal set of disjoint copies of J_i in $G[\pi_i]$, then there are subsets $\mathcal{J}'_i \subset \mathcal{J}_i$ such that $|\mathcal{J}'_i| \le \mu n^{1-\lambda}$, $i \in [\mathrm{wpn}(H)]$, and the only way to obtain a copy of J from sets $(J_1, J_2, ..., J_{\mathrm{wpn}(H)})$ where $J_i \in \mathcal{J}_i$, is if $J_i \in \mathcal{J}'_i$

for each $i \in [\mathrm{wpn}(H)]$. In other words, we are allowed to build a copy of J only by choosing its parts from the sets $\mathcal{J}_i'$.

For example, let Π be a partition as above and let G be such $G[\pi_i]$ is a clique for each $i \in [\mathrm{wpn}(H)]$. Let J be a path on $2\,\mathrm{wpn}(H)$ vertices, then J is decomposable into $\mathrm{wpn}(H)$ edges $J_1, J_2, ..., J_{\mathrm{wpn}(H)}$. In this example, for each $i \in [\mathrm{wpn}(H)]$, J_i is (Π, λ, i)-common for any $\lambda > 0$, let $\mathcal{J}_i$ be a maximum sets of disjoint edges in $G[\pi_i]$. The graph G is $(\Pi, 2, \lambda, \mu)$-exceptional if any induced $P_{2\,\mathrm{wpn}(H)}$ in G, which is build by choosing edges from $\mathcal{J}_i$, uses only edges in sets $\mathcal{J}_i' \subseteq \mathcal{J}_i$ for some sets $\mathcal{J}_i'$ such that $|\mathcal{J}_i'| \leq \mu|\mathcal{J}_i|$.

Let Π as above and let $h \in \mathbb{N}$, $\lambda \geq 0$ and $\mu \geq 0$, we define $\boldsymbol{\mathcal{C}(\Pi, h, \lambda, \mu)}$ to be the set of all (Π, h, λ, μ)-exceptional graphs. Note that in some cases we are interested in (Π, h, λ, μ)-exceptional graphs for $\mu = 0$, in this case we do not allow any copy of J as above.

Let $k \in \mathbb{N}$ and $\varepsilon > 0$ be the constants from Theorem 3.2.1 applied with $\mathrm{Forb}(H)$ and $\delta > 0$ sufficiently small. Let $\varepsilon' > 0$ be the constant from Theorem 3.2.3 applied with k.

Lemma 3.3.7. *Let* $\varepsilon, \varepsilon' > 0$ *as above,* $h \in \mathbb{N}$*,* $\lambda \in [0, \frac{\min\{\varepsilon,\varepsilon'\}}{8})$ *and* $\mu \in [0, \frac{1}{2})$*, then the number of graphs in* $\mathcal{C}(\Pi, h, \lambda, \mu)$ *is much smaller than the number of* Π*-good graphs.*

Proof. From Theorems 3.2.1 and 3.2.3 and similarly to the counting in the proof of Theorem 3.2.2, the number of possible graphs $G[\pi_i]$ for all $i \in [\mathrm{wpn}(H)]$ is at most

$$2^{n^{2-\varepsilon'}+2\delta\log\frac{1}{\delta}n^{2-\varepsilon}+(b+3)n} \leq 2^{n^{2-\varepsilon''/2}},$$

where $\varepsilon'' = \min\{\varepsilon, \varepsilon'\}$.

There are at most $(2^h \cdot 2^n)^{\mathrm{wpn}(H)}$ ways to choose the sets $\mathcal{J}_i$, $i \in [\mathrm{wpn}(H)]$. There are at most $(2^h \cdot 2^n)^{\mathrm{wpn}(H)}$ ways to choose the subsets $\mathcal{J}_i' \subseteq \mathcal{J}_i$, $i \in [\mathrm{wpn}(H)]$. There are at most $\left(\mu n^{1-\lambda}\right)!^{\mathrm{wpn}(H)-1}$ ways to partition the graphs in the sets $\mathcal{J}_i'$ into the different copies of J.

The number of possible edges between the graphs $J_1, J_2, ..., J_{\mathrm{wpn}(H)}$ is $m(J) = \prod_{1 \leq i < j \leq \mathrm{wpn}(H)} |V(J_i)| \cdot |V(J_j)| \leq \mathrm{wpn}(H)^2 \cdot h^2$, therefore the number of possible edge arrangements between the graphs $J_1, J_2, ..., J_{\mathrm{wpn}(H)}$ is $2^{m(J)}$. By our assumptions, we cannot have a copy of J by some choice of edges between graphs in $\mathcal{J}_i \setminus \mathcal{J}_i'$, therefore for every choice of sets $(J_1'', J_2'', ..., J_{\mathrm{wpn}(H)}'')$ where $J_i'' \in \mathcal{J}_i \setminus \mathcal{J}_i'$, $i \in [\mathrm{wpn}(H)]$, instead of $2^{m(J)}$

possible arrangements, there can be at most $2^{m(J)} - 1$. By Lemma 1.3.3, there are at least $\left(\frac{|\mathcal{J}_i| - |\mathcal{J}'_i|}{2}\right)^2 \geq \left(\frac{(1-\mu)n^{1-\lambda}}{2}\right)^2$ ways to choose the collections $(J''_1, J''_2, ..., J''_{\mathrm{wpn}(H)})$ where $J''_i \in \mathcal{J}_i \setminus \mathcal{J}'_i$, $i \in [\mathrm{wpn}(H)]$, such that no two collections intersect on more than one graph. Hence for each such choice of a collection, we have one less choice of arrangements between the sets. Therefore the number of graphs in $\mathcal{C}(\Pi, h, \lambda, \mu)$ is at most

$$\begin{aligned}
&2^m \cdot 2^{n^{2-\varepsilon''/2}} \cdot 2^{4\,\mathrm{wpn}(H)n} \cdot \left(\mu n^{1-\lambda}\right)!^{\mathrm{wpn}(H)-1} \cdot \left(\frac{2^{m(J)} - 1}{2^{m(J)}}\right)^{\left(\frac{(1-\mu)n^{1-\lambda}}{2}\right)^2} \\
&\leq 2^m \cdot 2^{2n^{2-\varepsilon''/2}} \cdot 2^{4\,\mathrm{wpn}(H)n} \cdot 2^{\mathrm{wpn}(H)n^{1-\lambda}\log n} \cdot 2^{-c(h,\mu)n^{2-2\lambda}}
\end{aligned}$$

where $c(H) = \log\left(\frac{2^{m(J)}}{2^{m(J)}-1}\right) \cdot \frac{(1-\mu)^2}{4}$. Using that $\lambda \leq \frac{\varepsilon''}{8}$, we get that the above is much smaller than the number of Π-good graphs. $\square$

3.3.3 The set $\mathcal{Y}(G)$

In this subsection we define a special set $Y(G) \subseteq V(G)$ in a graph G with respect to some parameters specified later. Roughly speaking, the vertices in $Y(G)$ are part of some constant sized sets of vertices with a non-typical behaviour. There are a few graphs G with a large set $Y(G)$. Moreover, it is possible to derive some useful properties of the graph $G[V(G) \setminus Y(G)]$.

Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(H)})$ be a ρ-almost equal partition of $[n]$ for some $\rho > 0$. Let $\lambda > 0$ and $i \in [\mathrm{wpn}(H)]$, a subgraph J of H is **(Π, λ, i)-linearly common** in G if there is a set $\mathcal{J}$ of disjoint copies of graphs isomorphic to J in $G[\pi_i]$ and $|\mathcal{J}| \geq \lambda n$. Note that if a subgraph J of H is (Π, λ, i)-linearly common, then it is also (Π, λ, i)-common.

Let $\mu > 0$, a set S such that $|S| \leq h$ and S for $j \neq i$, is **(Π, λ, i, μ)-linearly extremal** in G if there is a (Π, λ, i)-linearly common graph J and a graph J' which can be partitioned into $(G[S], J)$, but there are less than $\mu \cdot \lambda n$ vertex disjoint copies of J in $G[\pi_i]$ inducing J' with S.

Let G be a graph, we build the collection of disjoint subsets $\boldsymbol{\mathcal{Y}(\Pi, G, i)} := \boldsymbol{\mathcal{Y}\left(\Pi, G, \lambda, i, \frac{1}{2h^2+1}\right)}$ by adding greedily all the sets S which are $\left(\Pi, \lambda, i, \frac{1}{2h^2+1}\right)$-linearly extremal. Let $\boldsymbol{\mathcal{Y}(G)} := \cup_{i=1}^{\mathrm{wpn}(H)} \mathcal{Y}(\Pi, G, i)$, and $\boldsymbol{Y(G)} := \cup_{Y \in \mathcal{Y}(G)} Y$ and $\boldsymbol{y(G)} := |Y(G)|$.

Let Π be a partition as above, let $y \in [n]$ and let $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ be some sequence of families (not necessarily a $P(H)$-free sequence). Let $M(\Pi, \mathcal{F}, y)$ be equal to $\max_{Y \subset [n], |Y| \leq y} \prod_{i=1}^{\text{wpn}(H)} |(\mathcal{F}_i)_{|\pi_i \setminus Y|}|$. Note that the value $M(\Pi, \mathcal{F}, y)$ is similar to the value $F(H, \Pi)$ defined earlier, with the difference that in the case of $M(\Pi, \mathcal{F}, y)$ we remove some set of vertices Y such that $|Y| \leq y$ and do not restrict ourselves to $P(H)$-free sequences. Let $\lambda > 0$ and let $\boldsymbol{\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)}$ be the set of all graphs G which have properties (II) and (III) with respect to Π. Moreover, for each $i \in [\text{wpn}(H)]$, $G[\pi_i'] \in \mathcal{F}_i$ and $y(G) = y$.

Lemma 3.3.8. *Let H be a graph and let $\lambda > 0$. Let Π be a ρ-almost equal partition for $\rho = \rho(H, \xi)$ from Theorem 3.2.2 applied with H and $\xi > 0$ so $2\xi \log \frac{1}{\xi} < \log e \cdot \frac{\lambda}{2(2^{h^2+3} \cdot h \cdot \text{wpn}(H))}$. Let $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ be a sequence of families and let $y \in [n]$. The number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is at most*

$$2^{m(\Pi)} \cdot M(\Pi, \mathcal{F}, y) \cdot 2^{(b+3)n} \cdot 2^{-c(H,\lambda)ny},$$

where $c(H, \lambda) > 0$ is a constant which depends only on H and λ.

Proof. We bound from above the number of graphs which have properties (II)-(III) as in 3.2.2 and $y(G) = y$. There are at most $\binom{n}{b} \leq 2^n$ ways to choose the set $B \subset [n]$. There are at most 2^{bn} ways to choose the neighbourhoods of the vertices of B. There are at most 2^n ways to choose the set $Y(G)$ and there are at most y^y ways to partition it into the different subsets S. As mentioned we can assume that the graphs we count have property (III) of Theorem 3.2.2 with respect to Π. Using this property and Lemma 1.3.1, for any $i \in [\text{wpn}(H)]$ there are at most $b\binom{n}{\xi n} = b2^{2\xi \log \frac{1}{\xi} n}$ ways to choose the neighbourhood of any vertex $y \in \pi_i \cap Y$ with respect to $G[\pi_i]$. In total, there are at most $2^n \cdot y^y \cdot \left(b2^{2\xi \log \frac{1}{\xi} n}\right)^y$ possibilities for choosing $\mathcal{Y}$ and the neighbourhoods of each $y \in Y$ with respect to the part which it belongs to.

Assume that we fixed some choice of subgraphs on each of the π_i, $i \in [\text{wpn}(H)]$. We choose each edge between the parts independently at random with probability $\frac{1}{2}$.

Let $S \in \mathcal{Y}(G)$, by the definition of the set $\mathcal{Y}(G)$, the set S is $\left(\Pi, \lambda, i, \frac{1}{2^{h^2+1}}\right)$-linearly extremal and the following holds. Hence there is an index $i \neq j \in [\text{wpn}(H)]$ such that there is a (Π, λ, i)-linearly common graph J and a graph J' which can be partitioned into $(G[S], J)$, but there are less than $\frac{\lambda n}{2^{h^2+1}}$ vertex disjoint copies of J in $G[\pi_i]$ inducing J' with S. Let $\mathcal{J}$ be a maximum collection of disjoint copies of J in $G[\pi_i]$. The probability for any graph J' which can be partitioned into $(G[S], J)$ is $\frac{1}{2^{|S|\cdot|V(J)|}} \geq \frac{1}{2^{h^2}}$ where the inequality is due to the assumptions that $|S| \leq h$ and $|V(J)| \leq h$. The expected number of copies of each such graph J' is at least $\frac{|\mathcal{J}|}{2^{h^2}} \geq \frac{\lambda n}{2^{h^2}}$. Using Chernoff bound 1.3.6, the probability for S being $\left(\Pi, \lambda, i, \frac{1}{2^{h^2+1}}\right)$-linearly extremal is at most $e^{-\frac{\lambda n}{2^{h^2+3}}}$.

By the pigeonhole principle, there is an index $i \in [\mathrm{wpn}(H)]$, such that there is a collection $\mathcal{S}$ of at least $\frac{\mathcal{Y}}{\mathrm{wpn}(H)}$ sets from $\mathcal{Y}$ which are $\left(\Pi, \lambda, i, \frac{1}{2^{h^2+1}}\right)$-linearly extremal. The events that the sets in $\mathcal{S}$ are $\left(\Pi, \lambda, i, \frac{1}{2^{h^2+1}}\right)$-linearly extremal are independent. Hence the number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is at most

$$
\begin{aligned}
& 2^{m(\Pi)} \cdot M(\Pi, \mathcal{F}, y) \cdot 2^{(b+2)n} \cdot 2^n y^y \left(b 2^{2\xi \log \frac{1}{\xi} n}\right)^y \cdot \left(e^{-\frac{\lambda n}{2^{h^2+3}}}\right)^{\frac{\mathcal{Y}}{\mathrm{wpn}(H)}} \\
& \leq 2^{m(\Pi)} \cdot M(\Pi, \mathcal{F}, y) \cdot 2^{(b+3)n} \cdot 2^{(\log y + \log b + 2\xi \log \frac{1}{\xi} n) y} \cdot 2^{-\log e \cdot \frac{\lambda}{2^{h^2+3} \cdot h \cdot \mathrm{wpn}(H)} ny} \\
& \leq 2^{m(\Pi)} \cdot M(\Pi, \mathcal{F}, y) \cdot 2^{(b+3)n} \cdot 2^{(\log n + \log b) y - \log e \cdot \frac{\lambda}{2^{h^2+4} \cdot h \cdot \mathrm{wpn}(H)} ny},
\end{aligned}
$$

where the first inequality is due to the fact that every set $S \in \mathcal{Y}(G)$ is of size at most h, and therefore $y \leq h|\mathcal{Y}|$. The second inequality is due to the choice of ξ. We set $c(H, \lambda) = \log e \cdot \frac{\lambda}{2^{h^2+5} \cdot h \cdot \mathrm{wpn}(H)}$ and get the required inequality for n large enough. $\square$

We frequently apply the above bound in our arguments. As mentioned before we count the graphs G where for each $i \in [\mathrm{wpn}(H)]$, $G[\pi_i \setminus Y(G)] \in \mathcal{F}_i$ with respect to some sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ of families which is not necessarily $P(H)$-free. Usually we define this sequence by first defining some family of graphs $\mathcal{B}$ with some well specified properties on each π_i, $i \in [\mathrm{wpn}(H)]$, and then considering the sequence $\boldsymbol{\mathcal{F}(\mathcal{B})}$. The sequence $\mathcal{F}(\mathcal{B}) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ is defined by taking a maximal collection of graphs $\mathcal{F}_i$, $i \in [\mathrm{wpn}(H)]$, so for each choice of graphs $G_i \in \mathcal{F}_i$ we obtain a graph in $\mathcal{B}$.

3.3.4 Extendable subgraphs

In this subsection we describe some of the properties that we can derive about the subgraph $G[V(G)\backslash Y(G)]$ for G a Π-conformal graph and a set $Y(G)$ as defined in the previous subsection.

Let H be a graph and let $\mathcal{F}(H) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ be a $P(H)$-free sequence, let $k \in [\text{wpn}(H)]$ and let $\lambda > 0$. Similarly to before, a graph J is **λ-linearly common** in a graph G, if G contains at least $\lambda|V(G)|$ disjoint copies of J. A subgraph J of H is **$(\mathcal{F}(H), k, \lambda)$-extendable** if for any choice of graphs $G_i \in \mathcal{F}_i$, $i \in [\text{wpn}(H)] \setminus \{k\}$, there is a partition $(H_1, H_2, ..., H_k = J, ..., H_{\text{wpn}(H)}) \in P(H)$ where for each $j \in [\text{wpn}(H)] \setminus \{k\}$, H_j is λ-linearly common in G_j. A subgraph J of H is **λ-universally extendable** with respect to H if it is $(\mathcal{F}(H), k, \lambda)$-extendable for all $\mathcal{F}(H) \in \Phi(H)$ and $k \in [\text{wpn}(H)]$. Note that because in the definition of λ-universally extendable we require J to be $(\mathcal{F}(H), k, \lambda)$-extendable for all $\mathcal{F}(H) \in \Phi(H)$ then there is no need in fixing k.

For example, the graph P_4 is $\frac{1}{4}$-universally extendable with respect to all trees $T \in \mathcal{T}^{\text{npl}}$ (and actually we show later that P_4 is universally extendable in all trees with perfect matching). Let $T \in \mathcal{T}^{\text{npl}}$, as we show in Section 5.2.1, any tree $T \in \mathcal{T}^{\text{npl}}$ can be partitioned into P_6 and $\text{wpn}(T) - 2$ edges. The graph P_6 can be partitioned into P_4 and an edge and into P_4 and a non-edge. Hence there are partitions of T into P_4 and $\text{wpn}(T) - 1$ edges and into P_4, a non-edge and $\text{wpn}(T) - 2$ edges. By Theorems 1.1.12, 1.1.14, and 1.1.16 for a tree $T \in \mathcal{T}^{\text{npl}}$, for every $P(T)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ and for any choice of graphs $(G_1, G_2, ..., G_{\text{wpn}(H)})$ one of the following outcomes holds. Either each G_i, $i \in [\text{wpn}(T)]$, contains at least $\frac{1}{4}|V(G_i)|$ disjoint edges or one of the graphs, without loss of generality G_1, contains at least $\frac{1}{4}|V(G_1)|$ disjoint non-edges and the rest of the graphs G_i, $i \geq 2$, contain at least $\frac{1}{4}|V(G_i)|$ disjoint edges. Hence, for any choice of $k \in [\text{wpn}(T)]$, the rest of the graphs G_i, $i \in [\text{wpn}(T)] \setminus \{k\}$ either all contain many disjoint edges or one of the graphs contain many disjoint non-edges and the rest contain many disjoint edges. Therefore P_4 is $\frac{1}{4}$-universally extendable for trees $T \in \mathcal{T}^{\text{npl}}$.

Let $\mathcal{J} := \mathcal{J}(H)$ be the set of all graphs which are λ-universally extendable with respect to H and some $\lambda > 0$. Assume $\mathcal{J} \neq \emptyset$, and let $\lambda'(J)$ be the value with respect to which

$J \in \mathcal{J}$ is $\lambda'(J)$-universally extendable. Let $\lambda = \lambda(\mathcal{J}) = \min_{J \in \mathcal{J}} \lambda'(J)$. Let $\xi > 0$ be such $2\xi \log \frac{1}{\xi} < \log e \cdot \frac{\lambda}{2(2^{h^2+3} \cdot h \cdot \mathrm{wpn}(H))}$ and let $\rho > 0$ that we get from Theorem 3.2.2 applied with H and ξ as above. Let Π be a $\rho/4$-almost equal partition of $[n]$, for n large enough. Let $\boldsymbol{\mathcal{C}(\mathcal{J}, \Pi)}$ be the collection of Π-conformal graphs G such that there are $J \in \mathcal{J}$ and $i \in [\mathrm{wpn}(H)]$ such that J is an induced subgraph of $G[\pi_i \backslash Y(G)]$, where $Y(G)$ is the set of vertices as defined in Subsection 3.3.3.

Lemma 3.3.9. *Let H be a graph, $\mathcal{J} = \mathcal{J}(H)$, $\lambda = \lambda(\mathcal{J})$, $\xi > 0$, $\rho > 0$, and Π be a $\rho/4$-almost equal partition defined as above. The number of graphs in $\mathcal{C}(\mathcal{J}, \Pi)$ is much smaller than the number of Π-good graphs.*

Proof. Let $G \in \mathcal{C}(\mathcal{J}, \Pi)$ be a Π-conformal graph such that there are a λ-universally extendable $J \in \mathcal{J}$ and an index $i \in [\mathrm{wpn}(H)]$ such that $G[\pi_i \backslash Y(G)]$ contains a subgraph J' isomorphic to J. The graph G is Π-conformal, therefore by the definition of Π-conformal graph, it has property (I) from Theorem 3.2.2. Let $(G[\pi'_1], G[\pi'_2], ..., G[\pi'_{\mathrm{wpn}(H)}])$ be the $P(H)$-free partition which we get from property (I). By our assumptions the graph J is λ-universally extendable so therefore there is a partition $(H_1, H_2, ..., H_{i-1}, J, H_{i+1}, ..., H_{\mathrm{wpn}(H)})$ such that for each $j \in [\mathrm{wpn}(H)] \backslash \{i\}$, $G[\pi'_j]$ contains a set $\mathcal{H}_j$ of at least $\lambda |V(G[\pi'_i])| \geq \frac{\lambda n}{2\,\mathrm{wpn}(H)}$ disjoint copies of H_j.

By our assumptions $J' \notin \mathcal{Y}(G)$, so by the definition of $\mathcal{Y}(G)$, for each $j \in [\mathrm{wpn}(H)] \setminus \{i\}$, there is a subset $\mathcal{H}'_j \subset \mathcal{H}_j$ such that $|\mathcal{H}'_j| \geq \frac{\lambda n}{2^{h^2+1} \cdot 2\,\mathrm{wpn}(H)}$ and for each $H'_j \in \mathcal{H}'_j$, $G[J' \cup H'_j]$ is isomorphic to $H[J \cup H_j]$. Hence for any sequence $(H'_1, H'_2, ..., H'_{i-1}, H'_{i+1}, ..., H'_{\mathrm{wpn}(H)})$ such that $H'_j \in \mathcal{H}'_j$, $j \in [\mathrm{wpn}(H)] \setminus \{i\}$, there is an edge arrangement that cannot appear, otherwise we get an induced copy of H. Then by Lemma 3.3.7 applied with $h = |V(H)|$, $\lambda > 0$ sufficiently small and $\mu = 0$ (because we do not want to allow any copy of H), the number of such graphs is much smaller than the number of Π-good graphs. $\square$

Let $\mathcal{J}$, $\lambda = \lambda(\mathcal{J})$, $\xi > 0$, $\rho > 0$, Π be a $\rho/4$-almost equal partition as above. Let $f_{\mathcal{J}}(\ell)$ be such $|(\mathrm{Forb}(\mathcal{J}))_\ell| \leq 2^{f_{\mathcal{J}}(\ell)}$ for all $\ell \in \mathbb{N}$ large enough. Recall that $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is the set of all graphs G which have properties (II) and (III) with respect to Π. Moreover, for each $i \in [\mathrm{wpn}(H)]$, $G[\pi'_i] \in \mathcal{F}_i$ and $y(G) = y$.

Corollary 3.3.10. *Let H be a graph, $\mathcal{J} = \mathcal{J}(H)$, $\lambda = \lambda(\mathcal{J})$, $\xi > 0$, $\rho > 0$, and Π be $\rho/4$-almost equal partition. Let $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ such that $\mathcal{F}_i \subseteq \mathrm{Forb}(\mathcal{J})$ and let $f_{\mathcal{J}}(\ell)$ be as defined above.*

If $f_{\mathcal{J}}(n) = 2$, then there exists a constant $C(H) > 0$ which depends only on H such that if $y \geq C(H)$, then the number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is much smaller than the number of Π-good graphs.

If $f_{\mathcal{J}}(n) \geq 2$, then there exists a constant $C(H) > 0$ which depends only on H such that if $y \geq C(H) \cdot \frac{f_{\mathcal{J}}(n)}{n}$, then the number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is much smaller than the number of Π-good graphs.

Proof. We use the bound from Lemma 3.3.8 on the number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$. By our assumptions, $M(\Pi, \mathcal{F}, y) \leq 2^{f_{\mathcal{J}}(n)}$.

If $f_{\mathcal{J}}(n) = 2$, then the number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is at most

$$\begin{aligned} &2^{m(\Pi)} \cdot 2^{f_{\mathcal{J}}(n)} \cdot 2^{(b+3)n} \cdot 2^{-C(H) \cdot c(H,\lambda) n \cdot f_{\mathcal{J}}(n)} \\ &\leq 2^{m(\Pi)} \cdot 2^{(b+3)n} \cdot 2^{2(1-C(H) \cdot c(H,\lambda) n)}. \end{aligned}$$

If $f_{\mathcal{J}}(n) \geq 2$, then the number of graphs in $\mathcal{G}(\Pi, \mathcal{F}, y, \lambda)$ is at most

$$\begin{aligned} &2^{m(\Pi)} \cdot 2^{f_{\mathcal{J}}(n)} \cdot 2^{(b+3)n} \cdot 2^{-C(H) \cdot c(H,\lambda) n \cdot \frac{f_{\mathcal{J}}(n)}{n}} \\ &\leq 2^{m(\Pi)} \cdot 2^{(b+3)n} \cdot 2^{(1-C(H) \cdot c(H,\lambda)) f_{\mathcal{J}}(n)} \\ &\leq 2^{m(\Pi)} \cdot 2^{(b+3)n} \cdot 2^{(1-C(H) \cdot c(H,\lambda)) n}. \end{aligned}$$

In both case, if we set $C(H) \geq \frac{b+5}{c(H,\lambda)}$, we get a number which is much smaller than the number of Π-good graphs which is at least $2^{m(\Pi)}$. $\square$

We give one more definition related to the universally extendable definition. The following definition differs in two aspects. Firstly, if a graph J is λ-universally extendable then for every $P(H)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ and *for any* index $k \in [\mathrm{wpn}(H)]$, for any choice of graphs $G_i \in \mathcal{F}_i$, $[\mathrm{wpn}(H)] \backslash \{k\}$, it is possible to find many disjoint copies of a subgraph H_i in G_i where $(H_1, H_2, ..., H_{k-1}, J, H_{k+1} ..., H_{\mathrm{wpn}(H)}) \in P(H)$. On the other hand, in the following definition the graphs H_i do not necessarily exist for any choice of index $k \in [\mathrm{wpn}(H)]$, but there is some choice of such index. Moreover, in the following we allow to further partition the graph J into smaller subgraphs.

Let $\mathcal{F}(H) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence, let $\{i_1, i_2, ..., i_r\} \subseteq [\mathrm{wpn}(H)]$ be a set of different indices and let $\lambda > 0$. A subgraph J of H is $\boldsymbol{(\mathcal{F}(H), \{i_1, i_2, ..., i_r\}, \lambda)}$**-extendable**, if for any choice of graphs $G_j \in \mathcal{F}_j$, $j \in [\mathrm{wpn}(H)] \setminus \{i_1, i_2, ..., i_r\}$, there is a partition $(H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ such that the following is true. The graph J can be partitioned into $(H_{i_1}, H_{i_2}, ..., H_{i_r})$, and for each $j \in [\mathrm{wpn}(H)] \setminus \{i_1, i_2, ..., i_r\}$, H_j is λ-linearly common in G_j.

Let $\mathcal{F}(H) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence, let $\{i_1, i_2, ..., i_r\} \subseteq [\mathrm{wpn}(H)]$ be a set of different indices and let $\lambda > 0$. Let J be a $(\mathcal{F}(H), \{i_1, i_2, ..., i_r\}, \lambda)$-extendable graph. Let $\boldsymbol{\mathcal{C}(J, \mathcal{F}(H), \{i_1, i_2, ..., i_r\}, \Pi)}$ be the collection of Π-conformal graphs G such that there is a partition $(H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ where the graph J can be partitioned into $(H_{i_1}, H_{i_2}, ..., H_{i_r})$ and for each $i \in \{i_1, i_2, ..., i_r\}$, $G[\pi_i \setminus Y(G)]$ contains a graph H_i' isomorphic to H_i and $G[\cup_{i \in \{i_1, i_2, ..., i_r\}} V(H_i')]$ is isomorphic to J. Similarly to Lemma 3.3.9 it is possible to show the following lemma.

Lemma 3.3.11. *Let H be a graph, and let $\xi > 0$, $\rho > 0$, and Π be $\rho/4$-almost equal partition defined as above. Let $\mathcal{F}(H) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence, let $\{i_1, i_2, ..., i_r\} \subseteq [\mathrm{wpn}(H)]$ be a set of different indices and let $\lambda > 0$. Let J be a $(\mathcal{F}(H), \{i_1, i_2, ..., i_r\}, \lambda)$-extendable graph. The number of graphs in $\mathcal{C}(J, \mathcal{F}(H), \{i_1, i_2, ..., i_r\}, \Pi)$ is much smaller than the number of Π-good graphs.*

Let $r \in [\mathrm{wpn}(H)]$, a subgraph J of H $\boldsymbol{(r, \lambda)}$**-universally extendable** if for each $\mathcal{F}(H) \in \Phi(H)$, it is possible to find a set of r different indices $\{i_1, i_2, ..., i_r\} \subseteq [\mathrm{wpn}(H)]$, such that J is $(\mathcal{F}(H), \{i_1, i_2, ..., i_r\}, \lambda)$-extendable.

Let $1 \leq r \leq \mathrm{wpn}(H) - 1$, let $\mathcal{J}(r) := \mathcal{J}(H, r)$ be the set of all graphs which are (r, λ)-universally extendable with respect to H and assume $\mathcal{J} \neq \emptyset$. Let $\lambda'(r, J)$ be the value with respect to which $J \in \mathcal{J}(r)$ is $\lambda'(r, J)$-universally extendable. Let $\lambda(r) = \lambda(\mathcal{J}(r)) = \min_{J \in \mathcal{J}(r)} \lambda'(J)$ and let $\lambda = \min_{1 \leq r \leq \mathrm{wpn}(H)-1} \lambda(r)$. Let $\xi > 0$ be such $2\xi \log \frac{1}{\xi} < \log e \cdot$

$\frac{\lambda}{2(2^{h^2+3}\cdot h\cdot \mathrm{wpn}(H))}$ and let $\rho > 0$ that we get from Theorem 3.2.2 applied with H and ξ as above. Let $\boldsymbol{\mathcal{C}(\mathcal{J}(r), \Pi)}$ be the collection of Π-conformal graphs G such that there is $J \in \mathcal{J}(r)$ and a set of r different indices $\{i_1, i_2, ..., i_r\} \subseteq [\mathrm{wpn}(H)]$ where there is a partition $(H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ so the graph J can be partitioned into $(H_{i_1}, H_{i_2}, ..., H_{i_r})$ and for each $i \in \{i_1, i_2, ..., i_r\}$, $G[\pi_i \setminus Y(G)]$ contains a graph H'_i isomorphic to H_i and $G[\cup_{i\in\{i_1,i_2,...,i_r\}} V(H'_i)]$ is isomorphic to J. Also similarly to Lemma 3.3.9 it is possible to show the following lemma.

Lemma 3.3.12. *Let H be a graph, let $1 \leq r \leq \mathrm{wpn}(H) - 1$, $\mathcal{J}(r)$, $\lambda > 0$, $\xi > 0$, $\rho > 0$, and Π be $\rho/4$-almost equal partition defined as above. The number of graphs in $\mathcal{C}(\mathcal{J}(r), \Pi)$ is much smaller than the number of Π-good graphs.*

Finally, we give one more lemma which can be derived from 3.3.12. Let all the constants as chosen above. Let $f_1(n), f_2(n)$ be functions of n such that $\frac{f_{\mathcal{J}}(n)}{f_1(n)\cdot f_2(n)} = \mathrm{o}(1)$ for all $n \in \mathbb{N}$ and $f_1(n) \geq f_2(n)$. Let $\boldsymbol{\mathcal{C}(\mathcal{J}(r), \Pi, f_1(n), f_2(n))}$ be the set of graphs G such that there is a (r, λ)-universally extendable graph $J \in \mathcal{J}(r)$ which can be partitioned into $(H_1, H_2, ..., H_r)$, and the following is true for G. The graph G is Π-conformal and there are indices $\{i_1, i_2, ..., i_r\} \subseteq [\mathrm{wpn}(H)]$ such that $G[\pi_{i_j} \setminus Y(G)]$, for each $j \in [r-1]$, contains at least $f_1(n)$ disjoint copies of H_{i_j} and $G[\pi_{i_r} \setminus Y(G)]$ contains at least $f_2(n)$ disjoint copies of H_{i_r}.

Lemma 3.3.13. *Let H be a graph, let $2 \leq r \leq \mathrm{wpn}(H) - 1$, $\mathcal{J}(r)$, $\lambda > 0$, $\xi > 0$, $\rho > 0$, Π be a $\rho/4$-almost equal partition and $f_1(n), f_2(n)$ defined as above. If $f_{\mathcal{J}}(n) \geq n$ then the number of graphs in $\mathcal{C}(\mathcal{J}(r), \Pi, f_1(n), f_2(n))$ is much smaller than the number of Π-good graphs.*

Proof. Let $J \in \mathcal{J}(r)$ and let $(H_1, H_2, ..., H_r)$ be its partition. Let $\mathcal{H}_{i_j}$ be a maximum collection of disjoint copies of H_{i_j} in $G[\pi_{i_j} \setminus Y(G)]$. A graph G in which there are $H'_{i_j} \in \mathcal{H}_{i_j}$, $j \in [r]$, so $G[\cup_{j\in[r]} V(H'_{i_j})]$ is isomorphic to J is in $\mathcal{C}(\mathcal{J}(r), \Pi)$. By Lemma 3.3.12, the number of such graphs is much smaller than the number of Π-good graphs. Therefore we focus on graphs $G \in \mathcal{C}(\mathcal{J}(r), \Pi, f_1(n), f_2(n))$ where $G[\cup_{j\in[r]} V(H'_{i_j})]$ is not isomorphic to J for any choice of graphs $H'_{i_j} \in \mathcal{H}_{i_j}$, $j \in [r]$. Let $\mathcal{C}'$ be this set of graphs.

We use the bound from Lemma 3.3.8 on the number of graphs in $\mathcal{C}(\mathcal{J}(r), \Pi, f_1(n), f_2(n))$ with a set Y such that $|Y| = y$. By Lemma 3.3.9 and our assumptions $M(\Pi, \mathcal{F}, y) \leq 2^{f_{\mathcal{J}}(n)}$. Moreover, it is possible to find a collection $\mathcal{S}$ of sequences of graphs $(H'_{i_1}, H'_{i_2}, ..., H'_{i_r})$ so $H'_{i_j} \in \mathcal{H}_{i_j}$, $j \in [r]$, every two sequences in the collection intersect on at most one graph and $|\mathcal{S}| \geq \frac{f_1(n) \cdot f_2(n)}{4}$. As mentioned, for every sequence in $\mathcal{S}$ there is an edge arrangement that cannot appear. Hence the number of graphs in $\mathcal{C}'$ is at most

$$2^{m(\Pi)} \cdot 2^{f_{\mathcal{J}}(n)} \cdot 2^{(b+3)n} \cdot 2^{-c(H,\lambda)ny} \cdot 2^{-c(H)f_1(n) \cdot f_2(n)} = \mathrm{o}(2^{m(\Pi)}).$$

So in total we get that the number of graphs $\mathcal{C}(\mathcal{J}(r), \Pi, f_1(n), f_2(n))$ is much smaller than the number of Π-good graphs. $\square$

3.3.5 Ordinary sequences

The tools developed in Subsections 3.3.1-3.3.4 are used before the main body of the proofs in Section 5.2. The majority of the arguments in the proofs in Section 5.2 goes into showing that there is some sequence $\mathcal{F}' = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\mathrm{wpn}(H)})$ of families of graphs with some specific properties such that almost all Π-conformal graphs G have a partition $(G[\pi_1 \setminus W(G)], G[\pi_2 \setminus W(G)], ..., G[\pi_{\mathrm{wpn}(H)} \setminus W(G)])$ where $G[\pi_i \setminus W(G)] \in \mathcal{F}'_i$ for all $i \in [\mathrm{wpn}(H)]$ and $W(G) \subseteq V(G)$ is a very small set. Once we showed the existence of such sequence we apply the tool described in this subsection. That finishes the proof of the exact structure.

Let G be a graph and $\tau \in \left(0, \frac{1}{2}\right)$, a vertex v is **(τ, G)-trivial** if $|N_G(v)| \leq \tau n$ or $|\overline{N}_G(v)| \leq \tau n$.

Let $P = (H_1, H_2, ..., H_{\mathrm{wpn}(H)}) \in P(H)$ be a partition and let $Q = (G_1, G_2, ..., G_{\mathrm{wpn}(H)})$ be a sequence of graphs. Let $w' \notin \cup_{i \in [\mathrm{wpn}(H)]} V(G_i)$ and assume that it has some neighbourhood in $\cup_{i \in [\mathrm{wpn}(H)]} V(G_i)$. Let $i \in \mathrm{wpn}(H)$, we say that there is a **(P, w', i)-form** $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}'_i, ..., \mathcal{H}_{\mathrm{wpn}(H)})$ in Q if there is a vertex $w \in V(H_i)$ such that there is a set $\mathcal{H}'_i$ of disjoint copies of $H_i \setminus \{w\}$ in G_i, and for each $j \in [\mathrm{wpn}(H)] \setminus \{i\}$ there is a set $\mathcal{H}_j$ of disjoint copies of H_j in G_j. Moreover, w' is adjacent to each $H'_i \in \mathcal{H}'_i$ as w to $H_i \setminus \{w\}$ and w' is adjacent to each $H'_j \in \mathcal{H}_j$, $j \in [\mathrm{wpn}(H)] \setminus \{i\}$ as w to H_j. Note that if we have a sequence Q

and we want to choose edges between the graphs in the sequence without creating an induced copy of H, then there are edge arrangements that cannot appear between the graphs in the form.

Let $\tau \in \left(0, \frac{1}{2}\right), \lambda \geq 0, k \in \mathbb{N}$, a **$(\tau, \lambda, k)$-ordinary $P(H)$-free subsequence** is a sequence of families of graph $\mathcal{F}' = \mathcal{F}'(\tau, \lambda, k) = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\text{wpn}(H)})$ such that there is a $P(H)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(H)})$ and $\mathcal{F}'_i \subseteq \mathcal{F}_i$ for each $i \in [\text{wpn}(H)]$. Moreover, all sequences $Q = (G_1, G_2, ..., G_{\text{wpn}(H)})$ for any $G_i \in \mathcal{F}'_i$, $i \in [\text{wpn}(H)]$ has the following properties,

(a) For each $i \in [\text{wpn}(H)]$, every $v \in V(G_i)$ is (τ, G_i)-trivial.

(b) Let $w' \notin \cup_{i \in [\text{wpn}(H)]} V(G_i)$ and assume that it has some neighbourhood in $\cup_{i \in [\text{wpn}(H)]} V(G_i)$ so for each $i \in [\text{wpn}(H)]$, $G_i \cup \{w'\} \notin \mathcal{F}_i$. Let $\ell_i(w') = \min\{|N(w') \cap V(G_i)|, |\overline{N}(w') \cap V(G_i)|\}$, $i \in [\text{wpn}(H)]$, let $\ell(w') = \min_{i \in [\text{wpn}(H)]} \ell_i(w')$ and let $i \in [\text{wpn}(H)]$ be such that $\ell(w') = \ell_i(w')$. Assume also that for each $j \in [\text{wpn}(H)] \setminus \{i\}$, $\ell_j(w') \geq \tau n$, that is w' is not (τ, G_j)-trivial. Then there is (P, w', i)-form $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}'_i, ..., \mathcal{H}_{\text{wpn}(H)})$ in Q so there is $\gamma \in \left[0, \frac{\tau}{2^{4h^2}}\right)$ such that one of the following conditions holds.

  (1) $|\mathcal{H}'_i| \geq \frac{\ell(w') - \gamma n}{(\log n)^k}$ and $|\mathcal{H}_j| \geq \frac{\tau n}{4h}$ for each $j \in [\text{wpn}(H)] \setminus \{i\}$.

  (2) $|\mathcal{H}'_i| \geq \frac{\ell(w')}{(\log n)^k}$ and $|\mathcal{H}_j| \geq n^{1-\lambda}$ for each $j \in [\text{wpn}(H)] \setminus \{i\}$.

Let $\mathcal{F}' = \mathcal{F}'(\tau, \lambda, k) = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\text{wpn}(H)})$ be a (τ, λ, k)-ordinary $P(H)$-free subsequence for some $\tau \geq 0, \lambda \geq 0$ and $k \in \mathbb{N}$. Let Π be a ρ-almost equal partition for some $\rho > 0$, and let $k' \in \mathbb{N}$. Let $\boldsymbol{\mathcal{B}(\mathcal{F}', \Pi, k')}$ be the set of Π-bad graphs G such that the minimal subset of vertices $W \subseteq V(G)$ where for each $i \in [\text{wpn}(H)]$, $G[\pi_i \setminus W] \in \mathcal{F}'_i$ is of size $|W| \leq (\log n)^{k'}$. Note that for any Π-bad G, the set W is not empty, because otherwise it contradicts the definition of Π-bad graph.

Similarly to the earlier definition of a linearly extremal set we define a **(Π, i, λ, μ)-extremal** set S if there is a (Π, λ, i)-common graph J and a graph J' which can be partitioned into $(G[S], J)$, but there are less than $\mu|\mathcal{J}|$ vertex disjoint copies of J in π_i inducing J' with S where $\mathcal{J}$ is the collection of disjoint copies of J.

Let $\tau \in \left(0, \frac{1}{2}\right)$, $\lambda > 0$, $\mu > 0$ and $i \neq j \in [\text{wpn}(H)]$ and $|V(H)| = h$, a set S such that $|S| \leq h$ is **$(\boldsymbol{\tau, \lambda, \mu, \Pi, i, j})$-special** if every $s \in S$ is $(\tau, G[\pi_i])$-trivial and S is (Π, j, λ, μ)-extremal.

Let $K \in \mathbb{N}$ and $\varepsilon > 0$ be the constants from Theorem 3.2.1 applied with Forb(H) and $\delta > 0$ sufficiently small. Let $\varepsilon' > 0$ be the constant from Theorem 3.2.3 applied with K. Let $\lambda \in [0, \frac{\min\{\varepsilon, \varepsilon'\}}{8})$.

Let Π be a ρ-almost equal partition of $[n]$ for $\rho = \rho(H, \xi)$ from Theorem 3.2.2 applied with H and $\xi > 0$ so $2\xi \log \frac{1}{\xi} < \log e \cdot \frac{\lambda}{2(2^{h^2+3} \cdot h \cdot \text{wpn}(H))}$.

Lemma 3.3.14. *Let $\varepsilon, \varepsilon' > 0$ as above, $\lambda \in [0, \frac{\min\{\varepsilon, \varepsilon'\}}{8})$, $\tau \in \left(0, \frac{1}{2}\right)$ and $k \in \mathbb{N}$. Let $\mathcal{F}' = \mathcal{F}'(\tau, \lambda, k) = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\text{wpn}(H)})$ be a (τ, λ, k)-ordinary $P(H)$-free subsequence. Let Π partition as above. Then the number of graphs in $\mathcal{B}(\mathcal{F}', \Pi, k')$ is much smaller than the number of Π-good graphs.*

Proof. Let $\mu \in \left(0, \frac{1}{2^{h^2+1}}\right)$ and let $G \in \mathcal{B}(\mathcal{F}', \Pi, k')$. Let $\mathcal{X}$ be the collection of all $(\tau, \lambda, \mu, \Pi, i, j)$-special sets for any $i \neq j \in [\text{wpn}(H)]$. Let X be the union of all vertices in sets in $\mathcal{X}$. Let $W \subseteq V(G)$ be the minimal subset of vertices such that for each $i \in [\text{wpn}(H)]$, $G[\pi_i \setminus W] \in \mathcal{F}'_i$. By the choice of G, $|W| \leq (\log n)^{k'}$. Let $W' = W \setminus X$.

Assume that $W' \neq \emptyset$. Let $w \in W'$, then either (i) the vertex w is not $(\tau, G[\pi_i])$-trivial for all $i \in [\text{wpn}(H)]$ or (ii) there is $i \in [\text{wpn}(H)]$ such that w is $(\tau, G[\pi_i])$-trivial and for any set $S \subseteq V(G)$ such that $w \in S$, $|S| \leq h$ and each $s \in S$ is $(\tau, G[\pi_i])$-trivial, the set S is not (Π, j, λ, μ)-extremal for all $j \in [\text{wpn}(H)] \setminus \{i\}$.

In case (i), by property (b) of $\mathcal{F}'(H)$, we have a (P, w', i)-form $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}'_i, ..., \mathcal{H}_{\text{wpn}(H)})$ in $(G[\pi_1 \backslash W], G[\pi_2 \backslash W], ..., G[\pi_{\text{wpn}(H)} \backslash W])$ for some partition $P = (H_1, H_2, ..., H_{\text{wpn}(H)}) \in P(H)$. Moreover $|\mathcal{H}'_i| \geq \min\left\{\frac{\tau n - \gamma n}{(\log n)^k}, \frac{\ell_i}{(\log n)^k}\right\} \geq \frac{\tau n}{2(\log n)^k} \geq n^{1-\lambda}$ for n large enough, and $|\mathcal{H}_j| \geq n^{1-\lambda}$ for all $j \in [\text{wpn}(H)] \setminus \{i\}$. Then because we need to forbid some edge arrangement for every choice of graphs from the form, by Lemma 3.3.7, the number of graphs in case (i) is much smaller than the number of Π-good graphs.

In case (ii), then as in the previous case, by property (b) of $\mathcal{F}'(H)$, we have a (P, w', i)-form $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}'_i, ..., \mathcal{H}_{\text{wpn}(H)})$ in $(G[\pi_1 \setminus W], G[\pi_2 \setminus W], ..., G[\pi_{\text{wpn}(H)} \setminus W])$ for some partition

$P = (H_1, H_2, ..., H_{\text{wpn}(H)}) \in P(H)$. Let $S' \in \mathcal{H}'_i$ and set $S = S' \cup \{w\}$. By our assumptions S is not (Π, j, λ, μ)-extremal for all $j \in [\text{wpn}(H)] \setminus \{i\}$, therefore there are subsets $\mathcal{H}'_j \subseteq \mathcal{H}_j$, $j \in [\text{wpn}(H)] \setminus \{i\}$ so for each $H'_j \in \mathcal{H}'_j$, $G[S \cup H'_j]$ is isomorphic to $H[H_i \cup H_j]$ and $|\mathcal{H}'_j| \geq \mu n^{1-\lambda}$. For every sequence $(H'_1, H'_2, ..., H'_{i-1}, H'_{i+1}, H'_{\text{wpn}(H)})$ where $H'_i \in \mathcal{H}'_i$, there is an edge arrangement that cannot appear. Therefore again by Lemma 3.3.7, the number of graphs in case (i) is much smaller than the number of Π-good graphs. Hence we can assume that $W' = \emptyset$.

Next we consider vertices in $W \cap X$. Let $W_i \subset W \cap X$, $i \in [\text{wpn}(H)]$, be the set of all vertices w for which the value $\ell(w) = \ell_i(w)$. Let A_1 be the set of vertices in $W \cap X$ for each of which condition (1) of (b) holds. Let $A_2 = (W \cap X) \setminus A_1$. By the definition for any $w \in A_2$ condition (2) of (b) holds.

We treat differently the case where $A_1 \neq \emptyset$ and the case where $A_1 = \emptyset$. Firstly assume that $A_1 \neq \emptyset$. Assume that there is an index $i \in [\text{wpn}(H)]$ and a vertex $w \in A_1 \cap W_i$ such that $\ell(w) \geq 2\gamma n$, then as before we get a (P, w', i)-form $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}'_i, ..., \mathcal{H}_{\text{wpn}(H)})$ for some $P \in P(H)$ such that $|\mathcal{H}'_i| \geq \frac{\gamma n}{(\log n)^k}$ and also $|\mathcal{H}_j| \geq \frac{\tau n}{4h} \geq \frac{\gamma n}{(\log n)^k}$ for each $j \in [\text{wpn}(H)] \setminus \{i\}$. Again by Lemma 3.3.7, the number of such graphs is much smaller than the number of Π-good graphs. Therefore we assume that $\ell(w) \leq 2\gamma n$ for all $w \in A_1$.

Let $A'_2 \subseteq A_2$ be the set of vertices w such that $\ell(w) \geq h(\log n)^{k'}$. If $A'_2 \neq \emptyset$, let $\ell^* = \max_{w \in A'_2} \ell(w)$. From the definition we have that $\ell^* \geq \ell(w)$ for all $w \in A_2$.

Let $i \in [\text{wpn}(H)]$, the number of ways to choose the neighbourhood of a vertex $w \in A_1 \cap W_i$ with respect to π_i is at most $2^{2\gamma n + |W \cap X|}$. The number of ways to choose the neighbourhood of a vertex $w \in (A_2 \setminus A'_2) \cap W_i$ with respect to π_i is at most $2^{h(\log n)^{k'} + |W \cap X|}$. The number of ways to choose the neighbourhood of a vertex $w \in A'_2 \cap W_i$ with respect to π_i is at most $2^{\ell^* + |W \cap X|}$. Let $a_1 = |A_1|, a'_2 = |A'_2|, a''_2 = |A_2 \setminus A'_2|$. Therefore in total, the number of ways to choose the neighbourhoods of all vertices in $W \cap X$ with respect to the part that they are in, is at most $2^{2\gamma n a_1 + h(\log n)^{k'} a''_2 + \ell^* a'_2 + |W \cap X|^2}$.

Let $\mathcal{X}_1 \subseteq \mathcal{X}$ be the collection of all sets X such that $X \cap A_1 \neq \emptyset$. By the definition of $\mathcal{X}$, the sets in $\mathcal{X}_1$ are of size at most h, therefore there are at least $\frac{|A_1|}{\text{wpn}(H) \cdot h}$ disjoint sets in $\mathcal{X}_1$

which are contained in the same part of the partition Π. Let $S \in \mathcal{X}_1$ and assume without loss of generality that $S \subseteq \pi_i$ and let $j \in [\mathrm{wpn}(H)] \setminus \{i\}$ be the index such that S is $(\tau, \lambda, \mu, \Pi, i, j)$-special. Then there is a graph H_j which is (Π, λ, i)-common and S is $(\Pi, j, \frac{\tau}{h}, \mu)$-extremal. Hence the number of ways to choose the neighbourhood of S in π_j is at most $2^{(1-c(H)\frac{\tau\mu}{h})|S||\pi_i|}$ where $c(H) > 0$ is a constant which depends only on H.

Note that because $|A_1| \leq |W| \leq (\log n)^{k'}$ and every set $S \in \mathcal{X}$ has size at most h, we have that $|\cup_{X \in \mathcal{X}_1} X| \leq h(\log n)^{k'}$. Assume now also that $A_2' \neq \emptyset$ and let w be such that $\ell^* = \ell(w)$. Assume without loss of generality that $w \in W_i$, $i \in [\mathrm{wpn}(H)]$. By property (b) part (2), there is a (P, w, i)-form $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}_i', ..., \mathcal{H}_{\mathrm{wpn}(H)})$ for some $P \in P(H)$ such that $|\mathcal{H}_i'| \geq \frac{\ell^* - h(\log n)^{k'}}{(\log n)^k}$ and $|\mathcal{H}_j| \geq n^{1-\lambda}$ for each $j \in [\mathrm{wpn}(H)] \setminus \{i\}$. As in the proof of Lemma 3.3.8, we assume that we chose the subgraphs on the parts π_i, $i \in [\mathrm{wpn}(H)]$ and we choose every edge between the parts with probability $\frac{1}{2}$. Let $H_i' \in \mathcal{H}_i'$ and $j \in [\mathrm{wpn}(H)] \setminus \{i\}$, the probability for any edge arrangement between H_i' and a graph $H_j' \in \mathcal{H}_j$ is at least $\frac{1}{2^{h^2}}$. Hence the expected number of graphs from $\mathcal{H}_j$ which give the same edge arrangement with H_i' is at least $\frac{|\mathcal{H}_j|}{2^{h^2}} \geq \frac{n^{1-\lambda}}{2^{h^2}}$. We use Chernoff bound 1.3.6 to bound the probability that the number of such graphs is less than $\frac{n^{1-\lambda}}{2^{h^2+1}}$. We conclude that the number of graphs in $\mathcal{B}(\mathcal{F}', \Pi, k')$ in the case that $A_1 \neq \emptyset$ and $A_2' \neq \emptyset$ is at most

$$\begin{aligned}
&2^m \cdot F \cdot 2^{2\gamma n a_1 + h(\log n)^{k'} a_2'' + \ell^* a_2' + |W \cap X|^2 - \frac{c(H)\tau n a_1}{4h 2^{2h^2}} - \frac{\ell^* - h(\log n)^{k'}}{(\log n)^k} \cdot \frac{c'(H) n^{1-\lambda}}{2^{2h^2}}} \\
&\leq 2^m \cdot F \cdot 2^{h(\log n)^{2k'} + 2\gamma n a_1 - \frac{c(H)\tau n a_1}{4h 2^{2h^2}} + \ell^* (\log n)^{k'} - c'(H) \frac{\ell^* - h(\log n)^{k'}}{2^{h^2} (\log n)^k} \cdot n^{1-\lambda}} \\
&\leq 2^m \cdot F \cdot 2^{2h(\log n)^{2k'} - \frac{\tau n a_1}{c''(H)} + \frac{\ell^* - h(\log n)^{k'}}{2^{h^2} (\log n)^k} \cdot (2^{h^2} (\log n)^{k'+k} - c'(H) n^{1-\lambda})}.
\end{aligned}$$

The above bound is much smaller than $2^m \cdot F$ which is, from Observation 3.3.6, the lower bound on number of Π-good graphs.

In the above we showed that the number of graphs wuth $A_1 \neq \emptyset$ and $A_2' \neq \emptyset$ is much smaller than the number of Π-good graphs. Next we consider the case that $A_1 \neq \emptyset$ and

$A'_2 = \emptyset$. Similarly to before, the number of graphs in $\mathcal{B}(\mathcal{F}', \Pi, k')$ in this case is at most

$$\begin{aligned}
&2^m \cdot F \cdot 2^{2\gamma n a_1 + h(\log n)^{k'} a''_2 + |W \cap X|^2 - \frac{c(H)\tau n a_1}{4h2^{2h^2}}} \\
&\leq 2^m \cdot F \cdot 2^{2h(\log n)^{2k'} - \frac{\tau n a_1}{c(''H)}},
\end{aligned}$$

this is again much smaller than the lower bound of $2^m \cdot F$ on the number of Π-good graphs.

The last case we consider is if $A_1 = \emptyset$. In this case we use that each $w \in A_2$ has property (b) part (2). Let $\ell^* = \max_{w \in A_2} \ell(w)$ and let w be such $\ell(w) = \ell^*$, assume without loss of generality that $w \in W_i$, $i \in [\mathrm{wpn}(H)]$. Repeating the arguments as before we get that in this case the number of graphs in $\mathcal{B}(\mathcal{F}', \Pi, k')$ is at most

$$\begin{aligned}
&2^m \cdot F \cdot 2^{\ell^* |W \cap X| + |W \cap X|^2} \cdot 2^{-c'(H)\frac{\ell^*}{(\log n)^k} \cdot n^{1-\lambda}} \\
&\leq 2^m \cdot F \cdot 2^{(\log n)^{2k'} + \left((\log n)^{k'} - c'(H)\frac{n^{1-\lambda}}{(\log n)^k}\right)\ell^*}
\end{aligned}$$

because we consider Π-bad graphs, it must be the case that $\ell^* \geq 1$. Hence the above is much smaller than the lower bound of $2^m \cdot F$ on the number of Π-good graphs. $\square$

Chapter 4

Witnessing Partition Number of a Bipartite Graph

In this chapter, we give a formula for the witnessing partition number $\mathrm{wpn}(H)$ where H is a bipartite graph. Recall that $\nu(H)$ is the size of a maximum matching in H. Let $\boldsymbol{H(h_1, h_2, ...h_k)}$ be a graph which is a disjoint union of complete bipartite graphs K_{h_i,h_i}, $i \in [k]$.

Lemma 4.0.1. *Let H be a bipartite graph with $|V(H)| \geq 3$, if $H = H(h_1, h_2, ...h_k)$ for some collection of vlaues $h_i \in \mathbb{N}$, $i \in [k]$, then $\mathrm{wpn}(H) = |V(H)| - \nu(H) = \nu(H) = \sum_{i=1}^{k} h_i$. Otherwise, $\mathrm{wpn}(H) = |V(H)| - \nu(H) - 1$.*

Proof. Let M be a matching in H of maximum size. Let $U = \cup_{e \in M} v(e)$ and let $L = V(H) \setminus U$. We denote $|V(H)| = h$, $|M| = m$, $|L| = \ell$, note that $|U| = 2m$ and $h = 2m + \ell$. Therefore we can restate what we want to show as, $\mathrm{wpn}(H) = h - m - 1 = (2m + \ell) - m - 1 = m + \ell - 1$, in the case that H is not a disjoint union of complete bipartite graphs with a perfect matching, and otherwise $\mathrm{wpn}(H) = m$.

Firstly assume that H is not a disjoint union of complete bipartite graphs with a perfect matching. Note that $\mathrm{wpn}(H) \geq m + \ell - 1$. Indeed, H cannot be partitioned into $m + \ell - 1$ cliques (which are either edges or vertices in a bipartite graph) and 0 stable sets. Assume otherwise, then out of the $m + \ell - 1$ cliques, at most m are edges, and therefore $m + \ell - 1$

cliques can contain at most $2m+\ell-1$ vertices in their union. This contradicts the assumption that $h = 2m + \ell$.

Next we show that $\mathrm{wpn}(H) \leq m + \ell - 1$. Indeed, let $c, s \in \mathbb{N}$ such that $c + s = m + \ell$, we need to show that H can be partitioned into c cliques and s stable sets. If $s \geq 2$, then we are done because H is bipartite and therefore can be partitioned into two stable sets and any number of cliques (which can be for example singleton vertices). If $s = 0$, then we are also done because by our choice of m and ℓ, H can be partitioned into $m + \ell$ cliques.

We are left with the case where $s = 1$. Firstly assume that $\ell \geq 1$. Note that $H[L]$ is a stable set, because otherwise we could add another edge to M which contradicts M being the maximum matching. Therefore if we partition H into the edges in M and L then we get a partition of H into at most $m + \ell - 1$ cliques and 1 stable set.

Secondly we consider the case that $\ell = 0$, that is H has a perfect matching. Let A, B be the partition of $V(H)$ into two stable sets. By our assumption about H, there is a connected component C of H which is not a complete bipartite graph. Between the vertices of any two edges of the perfect matching in C (and in the whole graph H) there can be, besides the edges of the matching, at most two additional edges. If there are two edges of the matching $e_1 = \{v_1, u_1\} \in M$ and $e_2 = \{v_2, u_2\} \in M$ such that there is exactly one additional edge between the vertices $\{v_1, u_1, v_2, u_2\}$, then we are done due to the following. Assume without loss of generality that $v_1, v_2 \in A$ and $u_1, u_2 \in B$, moreover there is no edge between v_1 and u_2 and there is an edge between u_1 and v_2, see Figure 4.1. The stable set $\{v_1, u_2\}$ together with the set of edges $M \setminus \{\{v_1, u_1\}, \{v_2, u_2\}\} \cup \{v_1, u_2\}$, is a partition of H into a stable set and $m-1$ cliques, as required. Otherwise, between any two edges of the matching there are either 0 or 2 edges. Let C' be a auxiliary graph such that $V(C')$ are the edges of the perfect matching in C, and $E(C')$ are the set of pairs of matching edges $\{e_1, e_2\}$ such that additionally to the edges of the matching, there are 2 more edges between the vertices $V(e_1) \cup V(e_2)$ (so in total there are all the possible edges between $V(e_1) \cup V(e_2)$). The graph C' is connected, because C is connected. Moreover, because C is not a complete bipartite graph, there are two vertices in C' which do not have an edge between them. Let $d_1 = \{v_1, u_1\}, d_2 = \{v_2, u_2\}$

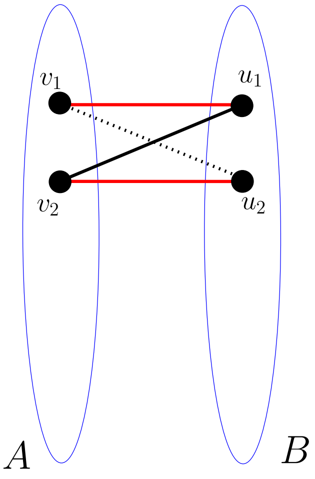


Figure 4.1: The edge arrangement between the vertices v_1, v_2, u_1, u_2.

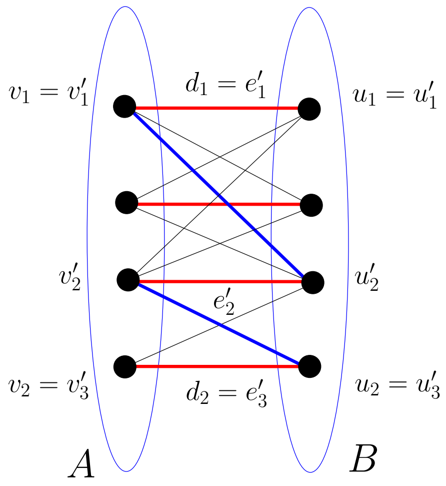


Figure 4.2: An example of P'.

be those two vertices. Because C' is connected, there is a path $P' = (d_1 = e'_1, e'_2, ..., d_2 = e'_k)$ between $d_1 = e'_1$ and $d_2 = e'_k$ in C'. Let $e'_i = \{v'_i, u'_i\}$, $i \in [k]$, without loss of generality, $v'_i \in A, u'_i \in B$, $i \in [k]$. See Figure 4.2 for an example of possible edge arrangement. Then $P = v_1, u'_2, v'_2, u'_3, v'_3..., v'_{k-1}, u_2$ is an alternating path in H. Therefore the stable set $\{u_1, v_2\}$ together with the set of edges $\Big(M \setminus \{e'_1, e'_2, ..., e'_k\}\Big) \cup \Big\{\{v'_1, u'_2\}, \{v'_2, u'_3\}..., \{v'_{k-1}, u'_k\}\Big\}$, is a partition of H into a stable set and $m-1$ edges, as required.

Now assume that $H = H(h_1, h_2, .., h_k)$ for some values $h_i \in \mathbb{N}$, $i \in [k]$. It is the case that $\text{wpn}(H) \geq m$ because H cannot be partitioned into 1 stable set and $m-1$ cliques. It is the case that $\text{wpn}(H) \leq m$, because similarly to before, it is possible partition H into c cliques and s stable sets such that $c + s = m + 1$ in the cases $s = 0$ or $s = 2$. In the case that $s = 1$, let $\{v, u\} \in M$, then $\{v\}$ together with $\{u\}$ and $M \setminus \{v, u\}$ is a partition of H into 1 stable set and m cliques. $\square$

Chapter 5

Obtaining Exact Partitions

In this chapter, we first reprove the result of Balogh and Butterfield [5] and show an exact structure theorem for critical graphs. We prove that graphs which are trees without a perfect matching are critical. The main part of this chapter is dedicated to the proof of an exact structure theorem for trees with a perfect matching.

5.1 Critical Graphs

We recall the definition that was given by Balogh and Butterfield in [5] of a critical graph. This definition was presented in Chapter 2. Let $\mathcal{F}(H, s, c)$ denote the set of minimal (by induced containment) graphs F such that H can be covered by s stable sets, c cliques, and F. A graph is critical if for all s, c such that $s + c = \mathrm{wpn}(H) - 1$ and large enough $n \in \mathbb{N}$, $|\left(\mathrm{Forb}(\mathcal{F}(H, s, c))\right)_n| \leq 2$, or equivalently by Lemma 2.2.10, $\left(\mathrm{Forb}(\mathcal{F}(H, s, c))\right)_n \subseteq \{S_n, K_n\}$.

As mentioned in Chapter 2, the graph C_4 is critical and every odd cycle of length at least 7 is critical. Let us consider two additional examples for a critical graph. The star T_3 with 4 vertices (3 leaves) is critical. It is not hard to check that $\mathrm{wpn}(S_3) = 2$. The set $\mathcal{F}(S_3, 1, 0)$ is a non-edge, therefore $(\mathrm{Forb}(\mathcal{F}(S_3, 1, 0)))_n$ is a clique for all $n \in \mathbb{N}$. The set $\mathcal{F}(S_3, 0, 1)$ is a vertex, therefore $(\mathrm{Forb}(\mathcal{F}(S_3, 0, 1)))_n$ is an empty set for all $n \in \mathbb{N}$. Another example for a critical graph is a path P_7 on 7 vertices. It is not hard to check that $\mathrm{wpn}(P_7) = 3$. The set

$\mathcal{F}(P_7, 2, 0)$ is $P_3, \overline{P_3}$, the set $\mathcal{F}(P_7, 1, 1)$ contains a non-edge and the set $\mathcal{F}(P_7, 0, 1)$ contains a vertex. In all the cases, the set $(\mathrm{Forb}(\mathcal{F}(P_7, c, s)))_n$ is either a clique or a stable set or both, or none.

We give a proof for Balogh and Butterfield's theorem using the results in [2] and [47] and also the tools from Chapter 3. Let $\mathcal{W}(H)$ be the collection of all pairs (s, c) such that H cannot be partitioned into s stable sets and c cliques where s, c are such that $s + c = \mathrm{wpn}(H)$. Let $\mathcal{Q}(H, s, c)$ be the set of all graphs that can be partitioned into s stable sets and c cliques. Let $\mathcal{Q}(H) = \cup_{(s,c) \in \mathcal{W}(H)} \mathcal{Q}(H, s, c)$.

Theorem (2.2.9, [5])**.** *Let H be a graph with $\mathrm{wpn}(H) \geq 2$. Almost all H-free graphs are in $\mathcal{Q}(H)$ if and only if H is critical.*

For the proof we need a structural lemma that was also shown in [5].

Lemma 5.1.1 ([5])**.** *If H is a critical graph, then for any s, c such that H cannot be partitioned into s stable sets and c cliques the following holds.*

- *If $s \geq 1$, then $\mathcal{F}(H, s-1, c)$ contains a graph that is the disjoint union of a clique and a vertex.*
- *If $c \geq 1$, then $\mathcal{F}(H, s, c-1)$ contains a graph that is the join of an stable set and a vertex.*

Proof of Theorem 2.2.9. We show that if H is critical then almost all H-free graphs are in $\mathcal{Q}(H)$. The other direction was explained in Chapter 2.

Let H be a critical graph and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence. First we argue that there is $\ell_0 \in \mathbb{N}$, such that for all $\ell \geq \ell_0$ and $i \in [\mathrm{wpn}(H)]$, $|(\mathcal{F}_i)_\ell| \leq 2$, that is by Lemma 2.2.10, $(\mathcal{F}_i)_\ell \subseteq \{K_\ell, S_\ell\}$. Indeed, by the definition of $P(H)$-free sequence, each of the families $\mathcal{F}_i$, $i \in [\mathrm{wpn}(H)]$ is infinite and hereditary so by Theorem 1.3.4, it must contain either all the stable sets or all the cliques or both. Let $i \in [\mathrm{wpn}(H)]$, let s be the number of families that contain all the stable sets and c be the number of families which contain all the cliques out of the families $\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{i-1}, \mathcal{F}_{i+1}, ..., \mathcal{F}_{\mathrm{wpn}(H)}$. Then $(\mathcal{F}_i)_\ell \subseteq$

$\left(\mathrm{Forb}(\mathcal{F}(H,s,c))\right)_\ell \subseteq \{S_\ell, K_\ell\}$ for any $\ell \geq h$ where $h = |V(H)|$, where the second containment is due to H being critical.

From the structure of the families in a $P(H)$-free sequence as shown above, we can conclude that any $P(H)$-free sequence is properly arranged and the families are stable. Hence from Theorem 3.3.5, to prove what is required, it is enough to show that the number of Π-bad graphs is much smaller than the number of Π-good graphs for any $\rho/4$-almost equal partition Π for $\rho > 0$ specified below.

Let $h = |V(H)|$ and let $\xi > 0$ be such that $2\xi \log \frac{1}{\xi} < \log e \cdot \frac{1}{2^{3 \cdot 2^{h^2+3} \cdot h \cdot \mathrm{wpn}(T)}}$. Let $\rho = \rho(H, \xi) > 0$ be the constant which we get from Theorem 3.2.2 for H and the above ξ. Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(T)})$ be a $\rho/4$-almost equal partition of $[n]$. We fix the partition Π for all of the following discussion. Let $n_i = |\pi_i|$, $i \in [\mathrm{wpn}(T)]$.

Let G be a Π-conformal graph. Let $\mathcal{Y}(\Pi, G, i) = \mathcal{Y}(\Pi, G, \frac{1}{4}, i, \frac{1}{2^{h^2+1}})$ be the collection of sets obtained by adding greedily sets S which are $(\Pi, \frac{1}{4}, i, \frac{1}{2^{h^2+1}})$-linearly extremal. Let $\mathcal{Y}(G) = \cup_{i=1}^{\mathrm{wpn}(T)} \mathcal{Y}(\Pi, G, i)$, $Y(G) = \cup_{Y \in \mathcal{Y}(G)} Y$ and $y(G) = |Y(G)|$ as defined in Chapter 3.

Let $\mathcal{F}(H) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence, then as argued before each of the families is either the family of all stable sets, all cliques, or both. Assume without loss of generality that families $\mathcal{F}_1, \mathcal{F}_2, .., \mathcal{F}_s$ contain all the stable sets and families $\mathcal{F}_{s+1}, \mathcal{F}_{s+2}, ..., \mathcal{F}_{\mathrm{wpn}(H)}$ contain all the cliques, let $c = \mathrm{wpn}(H) - s$. Let $F \in \mathcal{F}(H, s-1, c)$, then F is $(\mathcal{F}(H), \{i\}, \frac{1}{2h})$-extendable for all $1 \leq i \leq s$. Similarly, let $F \in \mathcal{F}(H, s, c-1)$, then F is $(\mathcal{F}(H), \{i\}, \frac{1}{2h})$-extendable for all $s+1 \leq i \leq \mathrm{wpn}(H)$. By Lemma 3.3.11, in almost all H-free graphs G, for each $1 \leq i \leq s$, $G[\pi_i \setminus Y(G)] \in (\mathrm{Forb}(\mathcal{F}(H, s-1, c)))_{|\pi_i \setminus Y(G)|} \subseteq \{S_{|\pi_i \setminus Y(G)|}, K_{|\pi_i \setminus Y(G)|}\}$ and for each $s+1 \leq i \leq c$, $G[\pi_i \setminus Y(G)] \in (\mathrm{Forb}(\mathcal{F}(H, s, c-1)))_{|\pi_i \setminus Y(G)|} \subseteq \{S_{|\pi_i \setminus Y(G)|}, K_{|\pi_i \setminus Y(G)|}\}$. From Lemma 3.3.8 and similarly to the proof of Lemma 3.3.10, there is a constant $C(H) > 0$, so the number of graphs G with a set $Y(G)$, so $y = |Y(G)| \geq C(H)$ is much more than the number of Π-good graphs.

Let $\mathcal{F}(H) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(H)})$ be a $P(H)$-free sequence (in this case we do not need to choose a subsequence). We show that there $\tau \in \left(0, \frac{1}{2}\right), \lambda > 0$ and $k \in \mathbb{N}$ for which $\mathcal{F}(H)$ is a (τ, λ, k)-ordinary $P(H)$-free subsequence. Let $i \in [\mathrm{wpn}(H)]$, and let $G \in \mathcal{F}_i$, then G is either a stable set or a clique, therefore for each $v \in V(G)$, either $|N_G(v)| = 0$ or $|\overline{N}_G(v)| = 0$, hence one of those values is at most τn for any $\tau \geq 0$, so property (a) holds for $\mathcal{F}(H)$.

Next we check property (b). Let $(G_1, G_2, .., G_{\text{wpn}(H)})$ be such that $G_i \in \mathcal{F}_i$ for each $i \in [\text{wpn}(H)]$. Let $w' \notin \cup_{i \in [\text{wpn}(H)]} V(G_i)$ such that for each $i \in [\text{wpn}(H)]$, $G_i \cup \{w'\} \notin \mathcal{F}_i$. Let $\ell_i(w') = \min\{|N(w') \cap V(G_i)|, |\overline{N}(w') \cap V(G_i)|\}$, $i \in [\text{wpn}(H)]$, let $\ell(w') = \min_{i \in [\text{wpn}(H)]} \ell_i(w')$ and let $i \in [\text{wpn}(H)]$ be such that $\ell(w') = \ell_i(w')$. By the definition we can also assume that $\ell_j(w') \geq \tau n$, that is $|N(w') \cap V(G_j)| \geq \tau n$ and $|\overline{N}(w') \cap V(G_j)| \geq \tau n$ for each $j \in [\text{wpn}(H)] \setminus \{i\}$. Assume without loss of generality that G_i is a clique, and out of the rest of the graphs G_j, $j \in [\text{wpn}(H)] \setminus \{i\}$, s are stable sets and $c - 1$ are cliques.

By Lemma 5.1.1, $\mathcal{F}(H, s, c-1)$ contains a graph H' which is a disjoint union of a clique and a vertex. Note that $|V(H')| \leq |V(H)| = h$. Let $P = (H_1, H_2, ..., H_{\text{wpn}(H)})$ be a partition of H into s stable sets, $c - 1$ cliques and the above mentioned graph $H_i = H'$. We argue that we can find a (P, w', i)-form $(\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}'_i, ... \mathcal{H}_{\text{wpn}(H)})$. Let w be the vertex in $V(H_i)$ which is not adjacent to any other vertex in H_i. Note that $\overline{N}(w') \cap V(G_i) \neq \emptyset$, as otherwise $G_i \cup \{w'\} \in \mathcal{F}_i$. Let $\mathcal{H}'_i$ be a partition of $\overline{N}(w') \cap V(G_i)$ into disjoint sets of size $|V(H_i)| - 1$. Note that $|\mathcal{H}'_i| \geq \frac{\ell(w)}{h}$. Let $j \in [\text{wpn}(H)] \setminus \{i\}$, and assume without loss of generality that G_j is a clique, because as mentioned $\ell_j(w') \geq \tau n$, it is possible to find in G_j a collection $\mathcal{H}_j$ of disjoint cliques of size $|V(H_j)| \leq h$ such that w' is as adjacent to each clique in $\mathcal{H}_j$ as w to H_j in $P(H)$. Moreover, $|\mathcal{H}_j| \geq \frac{\tau n}{2\,\text{wpn}(T) h}$. This shows that case (2) of property (b) holds for the form for any $\lambda > 0$ and $k \in \mathbb{N}$.

Finally, let G be a typical Π-conformal graph and let $W(G) = Y(G)$. Then as argued before $|W(G)| = |Y(G)| \leq C(H)$ for some constant $C(H)$ which depends only on H. Then by Lemma 3.3.14 the number of graphs with $W(G) \neq \emptyset$ is much smaller than the number of Π-good graphs. This completes the proof. □

In the following we focus on graphs H which are trees without a perfect matching. Recall that $\nu(T)$ is the size of maximum matching in T.

Theorem 5.1.2. *A tree T with $\text{wpn}(T) \geq 2$ and $2\nu(T) \leq |V(T)| - 1$ is a critical graph.*

In the proof of this theorem we use the following claims.

Claim 5.1.3. *Every tree T with $2\nu(T) \leq |V(T)| - 2$ can be partitioned into S_2 (a non-edge) and* $\operatorname{wpn}(T) - 1$ *edges or vertices.*

Proof. By Claim 4.0.1, $\operatorname{wpn}(T) = |V(T)| - \nu(T) - 1$. Let M be a perfect matching of maximum size in T. Let L be the set of vertices not covered by M. Let $m = |M|, \ell = |L|$, then $\operatorname{wpn}(T) = 2m + \ell - m - 1 = m + \ell - 1$.

By our assumption we consider trees T with $2\nu(T) \leq |V(T)| - 2$, therefore $|L| \geq 2$. Moreover, L is a stable set, because otherwise we could add an edge to the maximum matching. Let $\{v, u\} \subseteq L$, then $\{u, v\}$, $\ell - 2$ remaining singleton vertices in L and the m edges of the matching M give the required partition. Indeed the above partition is into $1 + \ell - 2 + m = m + \ell - 1 = \operatorname{wpn}(T)$ parts, as required. $\square$

Claim 5.1.4. *Let T be a tree such that $|V(T)| \geq 5$ and $2\nu(T) = |V(T)| - 1$, then T can be partitioned into*

- *P_3 and* $\operatorname{wpn}(T) - 1$ *edges, and*
- *$\overline{P_3}$ and* $\operatorname{wpn}(T) - 1$ *edges.*

Proof. Let M be a maximum matching in T and let v be the unique vertex which is not covered by M. If v is not a leaf, then we consider some alternating path P from v to one of the leaves ℓ' in T which is starting with an edge not in the matching and is ending with an edge in the matching. We can find such a path greedily. We define a new matching $M' := M \triangle P$, this is a valid matching and now the only vertex which is not matched is the leaf ℓ'. Therefore we can assume that v is a leaf.

Let u_1 be the unique neighbour of v, and let u_2 be the vertex such that $\{u_1, u_2\} \in M'$. Then $\{\ell, u_1, u_2\}$ induce a P_3 and $T' = T \backslash \{\ell, u_1, u_2\}$ has a perfect matching. This is a partition of T into P_3 and $\operatorname{wpn}(T) - 1$ edges. To show that we can partition T into $\overline{P_3}$ and $\operatorname{wpn}(T) - 1$ edges, we consider some edge $e \neq \{u_1, u_2\}$ such that $e \in M'$. We have that $e = \{u_1', u_2'\}$ together with ℓ induce a $\overline{P_3}$ and $F' = T' \setminus \{\ell, u_1', u_2'\}$ has a perfect matching. $\square$

Let $\mathcal{T}_{\text{star}}$ be the set of trees which are subdivided stars, that is each such tree is a star where every edge is subdivided exactly ones. Note that in each tree $T \in \mathcal{T}_{\text{star}}$ we have that $2\nu(T) = |V(T)| - 1$.

Claim 5.1.5. *Let T be a tree such that $|V(T)| \geq 5$ and $2\nu(T) = |V(T)| - 1$, then T can be partitioned into S_3 and $\operatorname{wpn}(T) - 1$ edges if and only if $T \notin \mathcal{T}_{\text{star}}$.*

Proof. Let T be a tree not in $\mathcal{T}_{\text{star}}$, and let M be a maximum matching such that the only vertex which is not covered by M is a leaf, we denote this vertex by ℓ. We can find such a matching as we did in the proof of Claim 5.1.4. Let $\{u_1, u_2\} \in M$ such that u_1 is a neighbor of ℓ. Let F be the tree (or a forest) that we get once we remove the vertices $\{\ell, u_1, u_2\}$. For a tree T such that $T \notin \mathcal{T}_{\text{star}}$, we have that there is at least one connected component with at least four vertices in F. In this component we can find an alternating path P between its ends which we denote as ℓ_s and ℓ_e. The path P starts and ends with an edge from the matching M. The set $\{\ell, \ell_s, \ell_e\}$ is stable. We define a new matching $M' = M \triangle P$, this matching shows that we can partition T into $S_3 = \{\ell, \ell_s, \ell_e\}$ and $\operatorname{wpn}(T) - 1$ edges.

Let $T \in \mathcal{T}_{\text{star}}$, then there are three types of vertices in such a tree, the centre c, the set D_1 of vertices at distance one from the centre, and the set D_2 of vertices at distance two from the centre. Assume to the contrary that it is possible to partition T into S_3 and $\operatorname{wpn}(T) - 1$ edges. If we choose for the stable set S_3 a vertex from D_1, then the leaf adjacent to it will not be matched in the remaining forest. Therefore we need to choose at least two vertices from D_2, but then only one vertex out of the two adjacent to them will be matched. Hence there is no such a stable set. □

Now we prove that every tree T with $2\nu(T) \leq |V(T)| - 1$ is critical.

Proof to Theorem 5.1.2. Let T with $2\nu(T) \leq |V(T)| - 2$. The tree T can be partitioned into two stable sets and $\operatorname{wpn}(T) - 2$ vertices. Therefore the set $\operatorname{Forb}(\mathcal{F}(T, s, \operatorname{wpn}(T) - s - 1))$ is an empty set for every $s \geq 2$. By Claim 5.1.3, $\mathcal{F}(T, 1, \operatorname{wpn}(T) - 2)$ contains an edge, therefore $\operatorname{Forb}(\mathcal{F}(T, 1, \operatorname{wpn}(T) - 2))$ contains only stable sets. Again by Claim 5.1.3, $\mathcal{F}(T, 0, \operatorname{wpn}(T) - 1)$ contains a non-edge, therefore $\operatorname{Forb}(\mathcal{F}(T, 0, \operatorname{wpn}(T) - 1))$ contains only cliques.

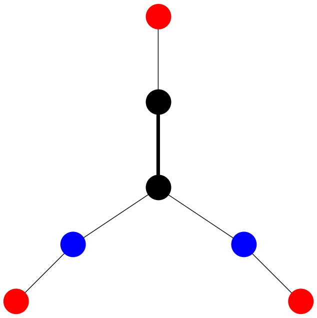


Figure 5.1: Partition of S_3^* into S_3, K_2 and S_2.

Let T with $2\nu(T) = |V(T)| - 1$. As before, the set $\text{Forb}(\mathcal{F}(T, s, \text{wpn}(T) - s - 1))$ is an empty set for every $s \geq 2$.

If $T \in \mathcal{T}_{\text{star}}$, if $T = P_5$, then $\text{wpn}(P_5) = 2$, and $\mathcal{F}(P_5, 1, 0)$ contains a non-edge. Otherwise, T can be partitioned into a graph S_3^* which is a star with leaves and with 3 subdivided edges and $\text{wpn}(T) - 3$ edges. The graph S_3^* can be partitioned into S_3, K_2 and S_2, see Figure 5.1. Therefore $\mathcal{F}(T, 1, \text{wpn}(T) - 2)$ contains a non-edge. In both cases, $\text{Forb}(\mathcal{F}(T, 1, \text{wpn}(T) - 2))$ contains only cliques. By Claim 5.1.4, $\mathcal{F}(T, 0, \text{wpn}(T) - 1)$ contains P_3 and $\overline{P_3}$, therefore $\text{Forb}(\mathcal{F}(T, 0, \text{wpn}(T) - 1))$ contains only cliques or stable sets.

If T is a tree with $2\nu(T) = |V(T)| - 1$ but not in $\mathcal{T}_{\text{star}}$, then $\mathcal{F}(T, 1, \text{wpn}(T) - c - 1)$ contains an edge by Claim 5.1.5. As for trees $T \in \mathcal{T}_{\text{star}}$, by Claim 5.1.4, $\mathcal{F}(T, 0, \text{wpn}(T) - 1)$ contains P_3 and $\overline{P_3}$, therefore $\text{Forb}(\mathcal{F}(T, 0, \text{wpn}(T) - 1))$ contains only cliques or stable sets. □

Next we present the proofs of the theorems regarding the structure of the families in the $P(T)$-free sequence for trees T with $2\nu(T) \leq |V(T)| - 1$. Those theorems were mentioned in the introduction.

Theorem (1.1.2). *Let T be a tree such that $2\nu(T) \leq |V(T)| - 2$, and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ be a $\mathcal{P}(T)$-free sequence, then for every $i \in [\text{wpn}(T)]$, $\mathcal{F}_i$ is the set of all cliques.*

Proof. First observe that no two families $\mathcal{F}_i$ and $\mathcal{F}_j$, $i \neq j$, can contain all the stable sets and this is because any tree T can be partitioned into two stable sets. Hence at most one family, without loss of generality $\mathcal{F}_1$, contains all the stable sets. Then each family $\mathcal{F}_i$, $i \in [\text{wpn}(T)]$, contains an edge. By Claim 5.1.3, the tree T with $2\nu(T) \leq |V(T)| - 2$ can be partitioned into a non-edge and $\mathrm{wpn}(T) - 1$ edges or vertices. Therefore $\mathcal{F}_1$ cannot contain a non-edge. Hence all the families are the families of all cliques. $\square$

Next we analyze the $\mathcal{P}(T)$-free sequence for trees T with $|V(T)| = 2\nu(T) - 1$. Naturally we need to separate between the trees with $2\nu(T) = |V(T)| - 1$ in $\mathcal{T}_{\text{star}}$ and such that are not in $\mathcal{T}_{\text{star}}$.

Theorem (1.1.4)**.** *Let T be a tree with $|V(T)| \geq 5$ such that $2\nu(T) = |V(T)| - 1$ and T is a subdivided star. Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then we have one of the following cases,*

- *(i) $\mathcal{F}_1$ is the set of all stable sets, and $\mathcal{F}_i$, for $2 \leq i \leq \mathrm{wpn}(T)$, is the set of all cliques.*
- *(ii) For every $i \in [\mathrm{wpn}(T)]$, $\mathcal{F}_i$ is the set of all cliques.*

Proof. Let us assume that one of the families, without loss of generality $\mathcal{F}_1$, contains all the stable sets. The tree T can be partitioned into two stable sets, so no other family $\mathcal{F}_i$, $i \geq 2$, contains all the stable sets, so it contains an edge.

Every tree $T \in \mathcal{T}_{\text{star}}$ can be partitioned into P_5 and $\mathrm{wpn}(T) - 2$ edges. The path P_5 can be partitioned into P_3 and S_2 and into $\overline{P_3}$ and S_2. Therefore because we assumed that $\mathcal{F}_1$ contains a non-edge, each $\mathcal{F}_i$ for $i \geq 2$ must be P_3 and $\overline{P_3}$-free. As mentioned each $\mathcal{F}_i$, $i \geq 2$ contains an edge so it must be the set of all cliques.

If $\mathcal{F}_1$ is not the set of all stable sets, then each $\mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$ contains an edge, and therefore by Claim 5.1.4, each $\mathcal{F}_i$ is a clique. $\square$

Theorem (1.1.5)**.** *Let T be a tree with $|V(T)| \geq 5$ such that $2\nu(T) = |V(T)| - 1$ and T is a subdivided star. Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then for every $i \in [\mathrm{wpn}(T)]$, $\mathcal{F}_i$ is the set of all cliques.*

Proof. As in the proof of Theorem 1.1.4 we know that only one family can contain all the stable sets, without loss of generality it is $\mathcal{F}_1$, therefore each of the families $\mathcal{F}_i$, $i \geq 2$, contains an edge. But by Claim 5.1.5, the family $\mathcal{F}_1$ can contain only a stable set of size at most 2, therefore it also contains an edge. Thus using Claim 5.1.4, we conclude that each family $\mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, is the family of all cliques. □

Proof of Theorems 1.1.3 and 1.1.6. The statements in the theorems are a direct corollary of Theorems 2.2.9, 5.1.2, 1.1.2, 1.1.4 and 1.1.5. □

5.2 Trees with a Perfect Matching

In this section we prove Conjecture 1.1.1 for trees with a perfect matching. That is, we prove Theorems 1.1.9,1.1.11,1.1.13,1.1.15, and 1.1.17. A first step in proving those theorems is finding the $\mathcal{P}(T)$-free sequences for each tree T with a perfect matching. Recall that $\mathcal{P}(T)$-free sequence is a sequence of maximal families $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ of graphs such that for any choice of graphs $G_i \in \mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, the resulting partition $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$ is $\mathcal{P}(T)$-free. A partition $P(G) = (G_1, G_2, ..., G_{\mathrm{wpn}(T)})$ of a graph G is $\mathcal{P}(T)$-free if for any partition $P(T) = (T_1, T_2, ..., T_{\mathrm{wpn}(T)})$ of T, there is an $i \in [\mathrm{wpn}(T)]$, such that T_i is not induced subgraph of G_i.

5.2.1 $P(T)$-free sequences for trees T with a perfect matching

As was mentioned earlier, we focus on trees with $\mathrm{wpn}(T) \geq 2$. Recall that $\mathcal{T}^{\mathrm{pl}}$ is the set of trees with a perfect matching where every non-leaf vertex has a neighbour which is a leaf. The set $\mathcal{T}^{\mathrm{pl}}_{\mathrm{star}} \subset \mathcal{T}^{\mathrm{pl}}$ is the family of all trees obtained from stars by subdividing every edge, except one, exactly once.

We denote by $\mathcal{T}^{\mathrm{npl}}$, the set of trees which have a perfect matching but not in $\mathcal{T}^{\mathrm{pl}}$. A tree T is in $\mathcal{S} \subset \mathcal{T}^{\mathrm{npl}}$ if there is some path P in T of length 6 or 8, such that the ends of P are leafs and the following property holds. Let $\mathcal{C}$ be the set of connected components in $T \setminus P$, then each components in $\mathcal{C}$ is an edge with the following additional properties. See example in Figure 5.2.

i. For $P = v_1, v_2, v_3, v_4, v_5, v_6$ of length 6, each of the components in $\mathcal{C}$ is joined by an edge to P at either v_3 or v_4. We denote this set as $\mathcal{S}^*_6$.

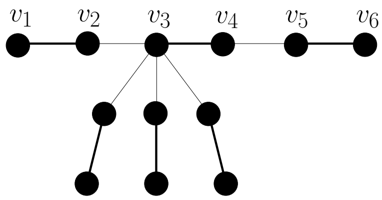


Figure 5.2: An example of a tree in $\mathcal{S}$; P_6 with three edges incident to v_3.

ii. For $P = v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8$ of length 8, each of the components in $\mathcal{C}$ is joined by an edge to P at either v_3 or v_6. We denote this set as $\mathcal{S}_8^*$.

Let $\mathcal{S} = \mathcal{S}_6^* \cup \mathcal{S}_8^*$. Let $\boldsymbol{\mathcal{T}_p}$ be the set of all trees which have a perfect matching and $\mathrm{wpn}(T) \geq 2$.

Claim 5.2.1. *Let $T \in \mathcal{T}_p$, then T has a unique perfect matching.*

Proof. Assume to the contrary that T has at least two different perfect matchings M and M'. Consider the graph $T' = (V(T), M \cup M')$, if $M \triangle M' \neq \emptyset$, T' will contain a cycle. But T' is a subgraph of T which is a tree and therefore does not contain cycles. Hence it must be that $M = M'$. $\square$

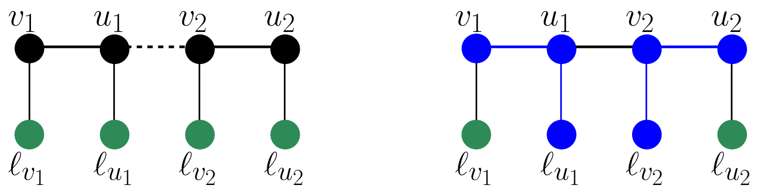
Let M_T be the unique perfect matching of $T \in \mathcal{T}_p$.

Lemma 5.2.2. *Let $T \in \mathcal{T}_p$, if T is not a path then M_T contains 3 edges e_1, e_2, e_3 such that $T[e_1 \cup e_2 \cup e_3]$ is isomorphic to P_3^*.*

Proof. If T is not a path then it has at least one vertex v such that $\deg(v) \geq 3$. Let v' be the neighbour of v such that $\{v, v'\} \in M_T$, and let $a \neq b$ be neighbours of v which are not v'. Due to the fact that T has a perfect matching, there are vertices a_1, b_1 such that $\{a, a_1\}, \{b, b_1\} \in M_T$. Hence the three edges $\{v, v'\}, \{a, a_1\}, \{b, b_1\}$ induce a graph isomorphic to P_3^*. $\square$

Corollary 5.2.3. *If $T \in \mathcal{T}^{\mathrm{pl}}$ then M_T contains 3 edges e_1, e_2, e_3 such that $T[e_1 \cup e_2 \cup e_3]$ is isomorphic to P_3^**

Let $T \in \mathcal{T}^{\mathrm{pl}}$, we denote by $\boldsymbol{T_s}$ the graph we get from T after removing all of its leaves.



(a) Partitions of P_4^* into S_4 and two edges.

(b) Partitions of P_4^* into $2P_3+S_2$.

Figure 5.3: Partitions of P_4^*.

Lemma 5.2.4. *If $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$, then T can be partitioned into P_4^* and* $\text{wpn}(T) - 3$ *edges. In particular, T can be partitioned into S_4 and* $\text{wpn}(T) - 1$ *edges, or $2P_3 + S_2$ and* $\text{wpn}(T) - 2$ *edges.*

Proof. Let $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$, then T_s contains a path of length at least 4, let $P = v_1, u_1, v_2, u_2$ be this path. Let $L = \{\ell_{v_1}, \ell_{u_1}, \ell_{v_2}, \ell_{u_2}\} \subset V(T)$ be the neighbouring leaves of v_1, u_1, v_2, u_2, respectively, those leaves exists due to the definition of trees in $\mathcal{T}^{\text{pl}}$. Then $\{v_1, u_1, v_2, u_2\} \cup L$ induces P_4^*.

The set L is a stable set and $M' = M_T \setminus \{\{v_1, \ell_{v_1}\}, \{u_1, \ell_{u_1}\}, \{v_2, \ell_{v_2}\}, \{u_2, \ell_{u_2}\}\} \cup \{\{v_1, u_1\}, \{v_2, u_2\}\}$ is a perfect matching in $T \setminus L$. See Figure 5.3a.

The sets $\{v_1, u_1, \ell_{u_1}\}, \{v_2, u_2, \ell_{v_2}\}$ induce two P_3, the remaining vertices $\{\ell_{u_1}, \ell_{u_2}\}$ induce a non-edge. See Figure 5.3b. ☐

As in the case of trees in $\mathcal{T}^{\text{pl}}$, we partition the tress in $\mathcal{T}^{\text{npl}}$ to the ones which can be partitioned into S_4 and $\text{wpn}(T) - 1$ edges and those which cannot. Before doing so we present a general claim about the trees in $\mathcal{T}^{\text{npl}}$.

Lemma 5.2.5. *Let $T \in \mathcal{T}^{\text{npl}}$, then M_T contains 3 edges e_1, e_2, e_3 such that $T[e_1 \cup e_2 \cup e_3]$ is isomorphic to P_6.*

Proof. Let $a_1 \in V(T)$ such that a_1 is not a leaf and it does not have a neighbour which a leaf, such a vertex exists by the definition of the trees in $\mathcal{T}^{\text{npl}}$. Let $m = \{a_1, a_2\} \in M_T$. By the choice of a_1, the vertices a_1 and a_2 have neighbours $b_1 \neq a_2$ and $b_2 \neq a_1$ respectively. Because

M_T is a perfect matching, there are vertices c_1 and c_2 such that $\{b_1, c_1\}, \{b_2, c_2\} \in M_T$. The three matching edges $\{a_1, a_2\}, \{b_1, c_1\}, \{b_2, c_2\}$ induce a graph isomorphic to P_6. $\square$

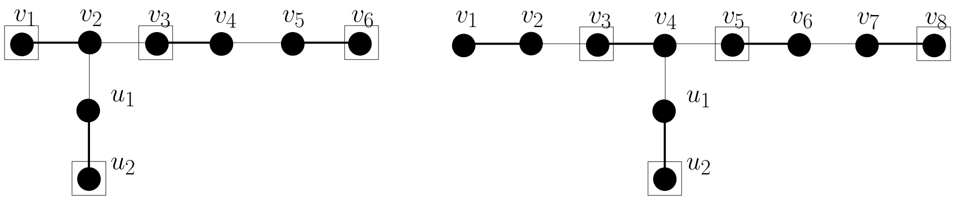
First we consider trees in $\mathcal{T}^{\text{npl}} \setminus \mathcal{S}$.

Lemma 5.2.6. *A tree* $T \in \mathcal{T}^{\text{npl}}$ *can be partitioned into a* S_4 *and* $\text{wpn}(T) - 1$ *edges if and only if* $T \in \mathcal{T}^{\text{npl}} \setminus \mathcal{S}$*.*

Proof. First we show that if a tree T is in $\mathcal{T}^{\text{npl}} \setminus \mathcal{S}$, then it can be partitioned into a S_4 and $\text{wpn}(T) - 1$ edges. Note that it is enough to find in T two induced disjoint alternating paths a_1, a_2, a_3, a_4 and b_1, b_2, b_3, b_4 ($\{a_1, a_2\}, \{a_3, a_4\}, \{b_1, b_2\}, \{b_3, b_4\} \in M_T$) such that a_1, a_4, b_1, b_4 is an induced stable set, because then this set can be partitioned into S_4 and 2 edges. Indeed, $M_T \setminus \{\{a_1, a_2\}, \{a_3, a_4\}, \{b_1, b_2\}, \{b_3, b_4\}\} \cup \{\{a_2, a_3\}, \{b_2, b_3\}\}$ is a perfect matching in $T \setminus \{a_1, a_4, b_1, b_4\}$. Hence our proof strategy is for any of the following cases to find two such alternating paths.

Let $T \in \mathcal{T}^{\text{npl}} \setminus \mathcal{S}$, and let $P = v_1, v_2, ...v_k$ be the longest alternating path in T, note that the ends of P are leafs and P is of even length. If $k \geq 10$, then the two paths will be v_1, v_2, v_3, v_4 and v_7, v_8, v_9, v_{10}. Otherwise, consider the connected components of $T \setminus P$, such components exist because of the following. By our choice of T it has $\text{wpn}(T) \geq 2$ and therefore $T \neq P_4$, and also $T \notin \mathcal{S}$ and $\{P_6, P_8\} \subset \mathcal{S}$ so $T \neq P_6$ and $T \neq P_8$. Note that each of those connected component is a subtree with a perfect matching, and therefore in particular contains at least 2 vertices, that is an edge. If a component contains more than 2 vertices, then it must contain at least 4 vertices and an alternating P_4 with the end edges in M_T. This P_4 together with the vertices v_1, v_2, v_3, v_4 gives the two needed paths. Now we are left only with the case that the remaining connected components are edges. We divide this case into some further cases:

If $P = P_6$, because $T \notin \mathcal{S}_6^*$, then $\deg(v_3) = \deg(v_4) = 2$ and either $\deg(v_2) \geq 3$ or $\deg(v_5) \geq 3$, assume that $\deg(v_2) \geq 3$ (the case that $\deg(v_5) \geq 3$ is symmetric). We denote by u_1 one of the vertices adjacent to v_2 and by u_2 the neighbor of u_1 which is not v_2, we have that $\{u_1, u_2\} \in M_T$, then the two paths will be v_1, v_2, u_1, u_2 and v_3, v_4, v_5, v_6. See Figure 5.4a for illustration.



(a) P_6 with $\deg(v_2) \geq 3$ (b) P_8 with $\deg(v_4) \geq 3$

Figure 5.4: Examples for the proof of claim 5.2.6.

If $P = P_8$, because $T \notin \mathcal{S}_8^*$, then $\deg(v_3) = \deg(v_6) = 2$, and the degree of either of the vertices v_2, v_4, v_5, v_7 is greater than two. In the case when $\deg(v_2) \geq 2$ (the case that $\deg(v_7) \geq 2$ is symmetric) we can find the two paths similarly to the case of $P = P_6$. If $\deg(v_4) \geq 2$ (the case that $\deg(v_5) \geq 2$ is symmetric) we denote by u_1 one of the vertices adjacent to v_4 and by u_2 its neighbour which is not v_4, we have $\{u_1, u_2\} \in M_T$, then the two paths can be v_3, v_4, u_1, u_2 and v_5, v_6, v_7, v_7. See Figure 5.4b for illustration.

Next we show that if T is in $\mathcal{S}$, then it cannot be partitioned into S_4 and $\text{wpn}(T) - 1$ edges. Note that if a tree T contains a vertex v such that at least two components of $T \setminus \{v\}$ are edges, then T can be partitioned into S_4 and $\text{wpn}(T) - 1$ edges if and only if the tree obtained from T by deleting one such component can. This can be seen by analyzing all the possible choices of two vertices, one from each such an edge, for the S_4. Let $T \in \mathcal{S}$, we show that T cannot be partitioned into S_4 and $\text{wpn}(T) - 1$ edges by induction on $|\mathcal{C}|$, where $\mathcal{C}$ is the collection of connected components in $T \setminus P$ and P is the longest alternating path in T. By the definition of $\mathcal{S}$ each component in $\mathcal{C}$ is an edge. If $|\mathcal{C}| = 1$, then $T \in \{P_6, P_8\}$ and we can see by inspection that neither can be partitioned into S_4 and $\text{wpn}(T) - 1$ edges. Otherwise, if $|\mathcal{C}| > 1$, then by the above observation, we can remove any such edge from the graph and proceed by induction. $\square$

Before we analyze the structure of the $\mathcal{P}(T)$-free sequence for trees $T \in \mathcal{T}_p$, we state a few observations. See Figures 5.5 and 5.6 for the illustration of the following observations.

Observation 5.2.7. *The graph P_3^* can be partitioned into two sets inducing each of the following pairs of graphs: $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, S_3), $(S_3, \overline{P}_3)$, (S_3, P_3) and also into (K_2, P_4),*

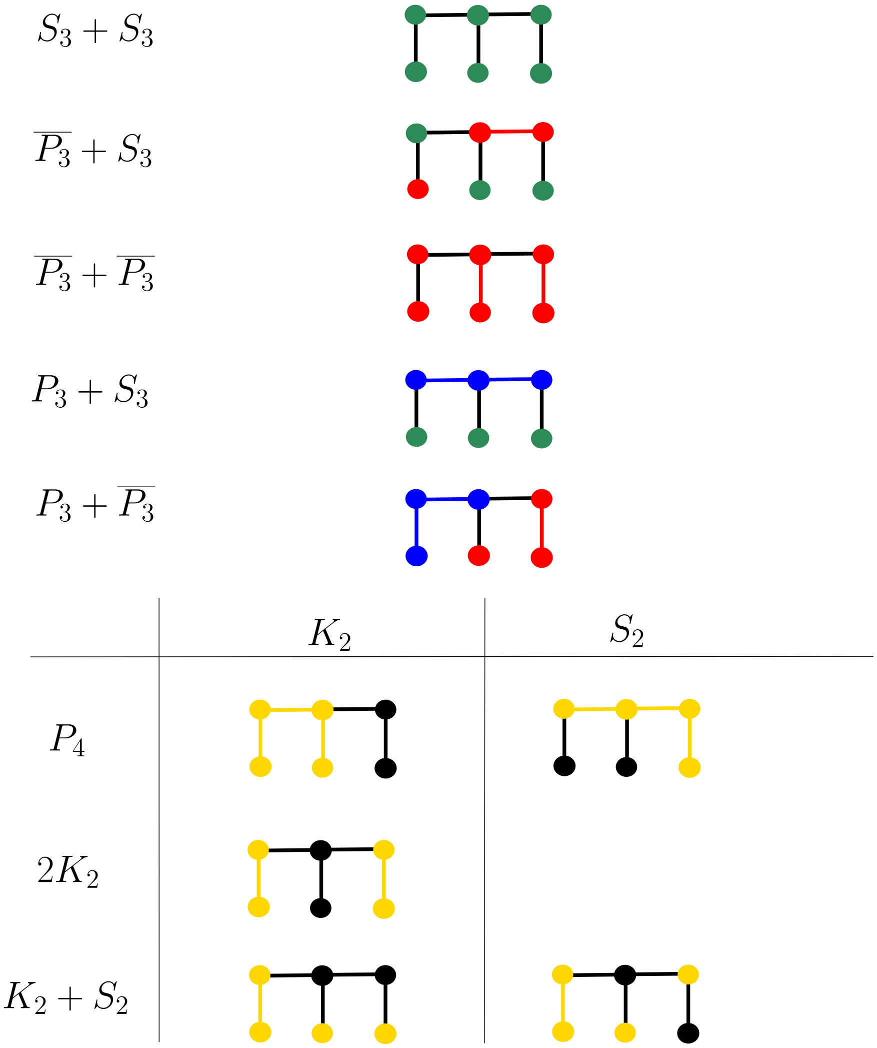
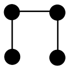


Figure 5.5: Partitions of P_3^*.

$(K_2, 2K_2)$, $(K_2, K_2 + S_2)$ and (S_2, P_4), $(S_2, K_2 + S_2)$.

Observation 5.2.8. *The graph P_6 can be partitioned into two sets inducing each of the following pairs of graphs: (P_3, P_3), $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, S_3),$(S_3, \overline{P}_3)$, and also into (K_2, P_4), $(K_2, 2K_2)$, $(K_2, P_3 + S_1)$ and (S_2, P_4), $(S_2, 2K_2)$, $(S_2, P_3 + S_1)$.*

Observation 5.2.9. *The graph P_8 can be partitioned into (S_3, P_3, K_2), $(S_2, 3K_2)$.*

We analyze the structure of a graph G without an induced $P_4 =$  and other graphs on 4 vertices.

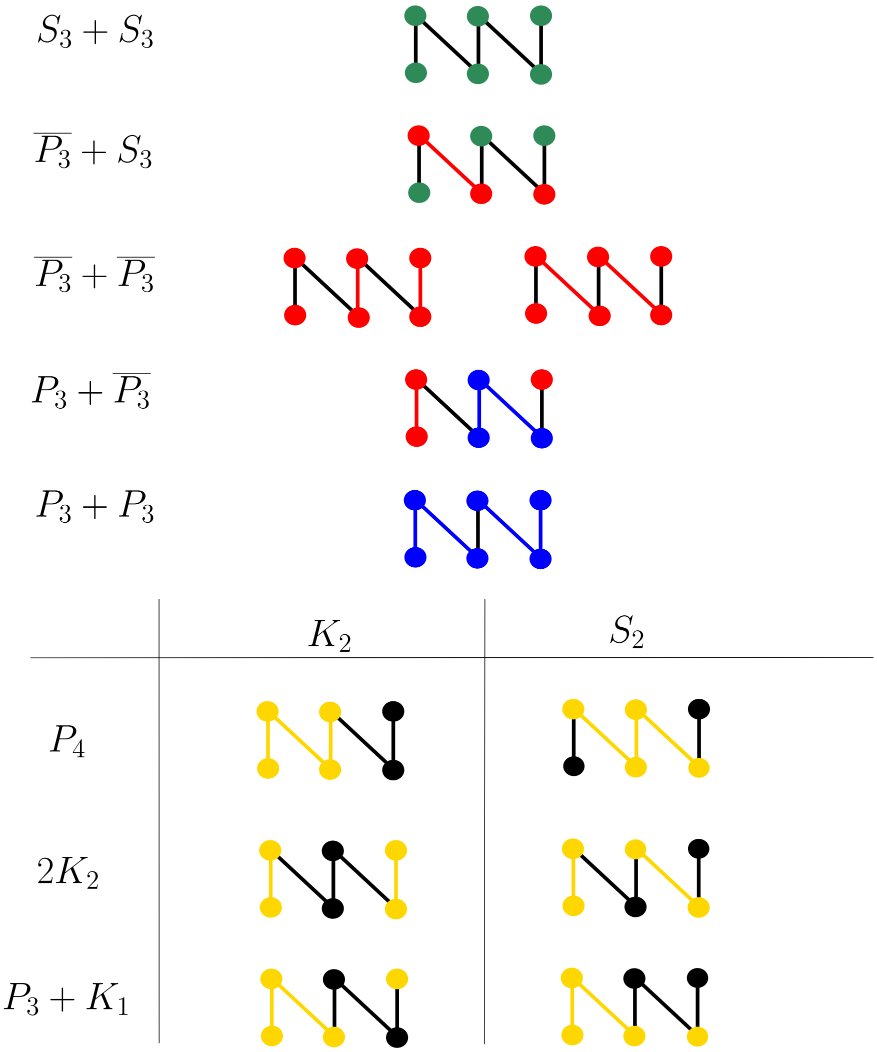


Figure 5.6: Partitions of P_6.

Lemma 5.2.10. *The following is a characterization of the families which are defined by forbidding the stated graphs.*

(i) $\mathrm{Forb}(\{P_4, 2K_2, K_2{+}S_2\}) = \mathcal{G}_2$*, that is the family of graphs which are the join of complete multi-partite graph with an isolated vertex.*

(ii) $\mathrm{Forb}(\{P_4, 2K_2, K_2 + S_2, S_4\}) = \mathcal{G}_3$*, that is the family of graphs which are the join of complete multi-partite graph with parts of size at most 2 and an isolated vertex.*

(iii) $\mathrm{Forb}(\{P_4, 2K_2, K_2 + S_2, P_3 + S_1\}) = \mathcal{G}_4$*, that is the family of graphs which are joins of graphs which are either stable set or a disjoint union of a vertex and a clique.*

(iv) $\mathrm{Forb}(\{P_4, 2K_2, K_2 + S_2, P_3 + S_1, S_4\}) = \mathcal{G}_5$*, that is the family of graphs which are joins of graphs which are either a stable set of size 3 or a disjoint union of a vertex and a clique.*

(v) $\mathrm{Forb}(\{P_4, 2K_2, P_3 + S_1\}) = \mathcal{G}_6$ *is the family of graphs which are joins of graphs which are a disjoint union of a clique and a stable set.*

Proof. If G does not contain a P_4 as an induced subgraph then either G or $\overline{G}$ is disconnected [53]. Therefore it is enough to analyze the structure of G when it is disconnected. Indeed, if it is not disconnected then it is a join of subgraphs which are disconnected (the connected components in $\overline{G}$). Hence we assume that G is disconnected and let $C_1, C_2, ..., C_t$ be the set of maximal connected components in G.

(i) If G does not have as an induced graphs $2K_2$ and $K_2 + S_2$, if $t \geq 3$, then G is a stable set. If $t = 2$, then one of the connected components is just a vertex, without loss of generality C_2, then C_1 does not contain a $\overline{P}_3$, therefore it is a complete multi-partite graph. All the following cases are the current case with additional restrictions, so we will just describe the additional structure due to them.

(ii) If G does not have as an induced graphs $2K_2$, $K_2 + S_2$ and S_4, with those restrictions $t \leq 3$, for $t = 3$ as before we get a S_3, for $t = 2$, we get that C_1 will be a multi-partite complete graph with parts of size at most 2.

(iii) If G does not have as an induced graphs $2K_2$, K_2+S_2 and P_3+S_1, for $t=2$, C_1 cannot contain either $\overline{P}_3$ or P_3, so it is a clique.

(iv) If G does not have as an induced graphs $2K_2$, K_2+S_2, P_3+S_1 and S_4, again with those restrictions $t \leq 3$, for $t=3$ as before we get a S_3 and for $t=2$ in this case C_1 is a clique again.

(v) If G does not have as an induced graphs $2K_2$ or P_3+S_1, then at most one of the connected components $C_1, C_2, ..., C_t$ can contain more than one vertex, assume it is C_1. Then C_1 must be P_3-free and because it is a connected component, it must be a clique.

This completes all the cases. □

Finally, we analyze the $\mathcal{P}(T)$-free sequence for trees $T \in \mathcal{T}_p$ and present some important lemmas for the next subsections. Let $\mathcal{G}_1 = \text{Forb}(\{S_3, \overline{P_3}\})$, that is all the graphs G, such that $\overline{G}$ is collection of edges and singleton vertices.

Theorem (1.1.8)**.** *Let T be a tree such that $T \in \mathcal{T}^{\text{pl}}$ and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then we have one of the following cases,*

(i) $\mathcal{F}_i = \mathcal{G}_1$ for each $i \in [\text{wpn}(T)]$.

(ii) The families can be reindexed such that $\mathcal{F}_1$ is the set $\mathcal{G}_2$ and the rest of the families are the sets of all cliques.

Proof. Note that by Corollary 5.2.3 and Observation 5.2.7, every tree $T \in \mathcal{T}^{\text{pl}}_{\text{star}}$ can be partitioned into,

(a) $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, P_3), $(S_3, \overline{P}_3)$, (S_3, S_3) and $\text{wpn}(T) - 2$ edges.

(b) P_4, $2K_2$, $K_2 + S_2$ and $\text{wpn}(T) - 1$ edges.

Firstly, assume that one of the families, without loss of generality, $\mathcal{F}_1$ contains the set of all the stable sets. The tree T can be partitioned into two stable sets of size at most $|V(T)| - 1$, so no other family $\mathcal{F}_i$, $i \geq 2$, can contain all the stable sets, so it contains an edge. Because

$\mathcal{F}_1$ contains an S_3, using (a) no other family can contain $P_3, \overline{P}_3$, so every other family is the set of all cliques.

Secondly, assume that no family contains all the stable sets. Therefore each family $\mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, contains an edge. Assume that there are at least two families which are not the sets of all the cliques. Then no family can contain either S_3 or $\overline{P}_3$, because then no other family can contain either one of $P_3, \overline{P}_3$, so they will be the families of all the cliques. Hence every family is the set of all complete multi-partite graphs with parts at most 2, that is equal to $\mathcal{G}_1$. Now, assume that there is at most one family $\mathcal{F}_1$ which is not the set of all cliques. Then by (b), the graphs in $\mathcal{F}_1$ are H-free for each $H \in \{P_4, 2K_2, K_2+S_2\}$, and therefore using Claim 5.2.10 $\mathcal{F}_1 = \mathcal{G}_2$. □

Theorem (1.1.10)**.** *Let T be a tree such that $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then we have one of the following cases,*

(i) The families can be reindexed such that $\mathcal{F}_1, \mathcal{F}_2$ are the set $\mathcal{G}_1$, and the rest of the families are the sets of all the cliques.

(ii) The families can be reindexed such that $\mathcal{F}_1$ is the set $\mathcal{G}_3$ and the rest of the families are the sets of all the cliques.

Proof. As in the previous theorem, using Corollary 5.2.3 and Observation 5.2.7, the trees in $\mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, can be partitioned into,

(a) $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, P_3), $(S_3, \overline{P}_3)$, (S_3, S_3) and $\mathrm{wpn}(T) - 2$ edges.

(b) P_4, $2K_2$, $K_2 + S_2$ and $\mathrm{wpn}(T) - 1$ edges.

Moreover, using Claim 5.2.4, the trees in $\mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, can be partitioned into,

(c) S_4 and $\mathrm{wpn}(T) - 1$ edges.

(d) $2P_3 + S_2$ and $\mathrm{wpn}(T) - 3$ edges.

Firstly, no family can contain all the cliques. Indeed, assume that one of the families, without loss of generality, $\mathcal{F}_1$ contains the set of all the stable sets. The tree T can be partitioned into two stable sets of size at most $|V(T)| - 1$, so no other family $\mathcal{F}_i$, $i \geq 2$, can contain all the stable sets, so it contains an edge. Hence by (c), $\mathcal{F}_1$ cannot contain S_4 and therefore it does not contain all the stable sets.

If there are two families, without loss of generality, $\mathcal{F}_1$ and $\mathcal{F}_2$ which contain P_3, then by (d) and the fact that every family contains an edge, every other family $\mathcal{F}_i$, $i \geq 3$, is the family of all cliques.

Finally, if there is at most one family $\mathcal{F}_1$ which is not the set of all cliques. Then by (a), $\mathcal{F}_1$ is H-free for each $H \in \{P_4, 2K_2, K_2{+}S_2, S_4\}$, and therefore using Claim 5.2.10 $\mathcal{F}_1 = \mathcal{G}_3$. $\square$

The next observation and a few lemmas are useful in the proofs of the exact structure of almost all T-free graphs for trees $T \in \mathcal{T}^{\mathrm{pl}}$ in Subsection 5.2.2. The observation and lemmas use the definitions from Section 3.3.

Observation 5.2.11. *Let $T \in \mathcal{T}^{\mathrm{pl}}$, then each $P(T)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ is properly arranged. Moreover, each of the families $\mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, is stable.*

Lemma 5.2.12. *Each of the graphs in $\{P_4, K_2 + S_2\}$ is $\frac{1}{4}$-universally extendable for all $T \in \mathcal{T}^{\mathrm{pl}}$. Moreover, each of the graphs in $\{2K_2, S_4\}$ is $\frac{1}{4}$-universally extendable for all trees $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$.*

Proof. Let $T \in \mathcal{T}^{\mathrm{pl}}$, then by Corollary 5.2.3 and Observation 5.2.7, T can be partitioned into (K_2, P_4), $(K_2, K_2 + S_2)$ and (S_2, P_4), $(S_2, K_2 + S_2)$ and $\mathrm{wpn}(T) - 2$ edges. By Theorem 1.1.8, for any $k \in [\mathrm{wpn}(T)]$, $\mathcal{P}(T)$-free sequence $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$, and $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$, where $G_i \in \mathcal{F}_i$ for all $i \in [\mathrm{wpn}(T)] \setminus \{k\}$, either each G_i contains $\frac{|V(G_i)|}{4}$ disjoint edges, or one of those graphs contains $\frac{|V(G_i)|}{4}$ disjoint non-edges and there contain $\frac{|V(G_i)|}{4}$ disjoint edges. Hence $\{P_4, K_2 + S_2\}$ is $\frac{1}{4}$-universally extendable for such tree T.

Let $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, then by Corollary 5.2.3, Observation 5.2.7 and Lemma 5.2.4, T can be partitioned into $2K_2$ or S_4 and $\mathrm{wpn}(T) - 1$ edges. By Theorem 1.1.10, for any $k \in [\mathrm{wpn}(T)]$, $\mathcal{P}(T)$-free sequence $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$, and $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$, where $G_i \in \mathcal{F}_i$ for all

$i \in [\mathrm{wpn}(T)] \setminus \{k\}$, each G_i contains $\frac{|V(G_i)|}{4}$ disjoint edges. Hence $\{2K_2, S_4\}$ is $\frac{1}{4}$-universally extendable for such tree T. $\square$

Lemma 5.2.13. *The graph P_3^* is $(2, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\mathrm{pl}}$.*

Proof. Let $T \in \mathcal{T}^{\mathrm{pl}}$, then by Corollary 5.2.3, T can be partitioned into P_3^* and $\mathrm{wpn}(T) - 2$ edges. By Theorem 1.1.8, for any $\mathcal{P}(T)$-free sequence $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$, there is a choice of $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, such that for any $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$, where $G_i \in \mathcal{F}_i$ for all $i \in [\mathrm{wpn}(T)] \setminus \{i_1, i_2\}$, each G_i contains $\frac{|V(G_i)|}{4}$ disjoint edges. Hence P_3^* is $(2, \frac{1}{4})$-universally extendable for such a tree T. $\square$

Lemma 5.2.14. *The graph P_4^* is $(3, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$.*

Proof. Let $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, by Lemma 5.2.4, T can be partitioned into P_4^* and $\mathrm{wpn}(T) - 3$ edges. By Theorem 1.1.10, for any $\mathcal{P}(T)$-free sequence $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$, there is a choice of different indices $\{i_1, i_2, i_3\} \in [\mathrm{wpn}(T)]$ (actually in this case, any choice of different indices will suffice), such that for any $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$, where $G_i \in \mathcal{F}_i$ for all $i \in [\mathrm{wpn}(T)] \setminus \{i_1, i_2, i_3\}$, each G_i contains $\frac{|V(G_i)|}{4}$ disjoint edges. Hence P_4^* is $(3, \frac{1}{4})$-universally extendable for such tree T. $\square$

Next we consider the trees in $\mathcal{T}^{\mathrm{npl}}$.

Theorem (1.1.12). *For a tree T such that $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$, and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then the families can be reindexed such that $\mathcal{F}_1$ is $\mathcal{G}_4$, and the rest of the families are the sets of all cliques.*

Proof. By Lemmas 5.2.5 and 5.2.2, each tree $T \in \mathcal{T}^{\mathrm{npl}}$ can be partitioned into P_6 and $\mathrm{wpn}(T) - 2$ edges or in the case it is not a path it can be partitioned into P_3^* and $\mathrm{wpn}(T) - 2$ edges. Hence by Observations 5.2.8 and 5.2.7 and 5.2.9, each tree $T \in \mathcal{T}^{\mathrm{npl}} \cap \mathcal{S}$ can be partitioned into,

(a) (P_3, P_3), $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, S_3), $(S_3, \overline{P}_3)$, (S_3, P_3) and $\mathrm{wpn}(T) - 2$ edges.

(b) P_4, $2K_2$, $K_2 + S_2$, $P_3 + K_1$ and $\mathrm{wpn}(T) - 1$ edges.

Assume that there is a family, without loss of generality, $\mathcal{F}_1$, which contains either of $\{S_3, P_3, \overline{P_3}\}$, then by (a) the rest of the families cannot contain either of the graphs from $\{S_3, P_3, \overline{P_3}\}$ and therefore are the sets of all cliques. Then by (b) and Claim 5.2.10 $\mathcal{F}_1 = \mathcal{G}_4$.

If no family contains either of graphs in $\{S_3, P_3, \overline{P_3}\}$, then all the families are the sets of all cliques. □

Theorem (1.1.14)**.** *Let $T = P_6$ and let $(\mathcal{F}_1, \mathcal{F}_2)$ be the $P(P_6)$-free sequence, then we have the following cases,*

(i) The families can be reindexed such that $\mathcal{F}_1$ is the family of all stable sets and $\mathcal{F}_2 = \mathcal{G}_1$.

(ii) The families can be reindexed such that $\mathcal{F}_1$ is the family of all cliques and $\mathcal{F}_2 = \mathcal{G}_6$.

Proof. By Observation 5.2.8, P_6 can be partitioned into,

(a) (P_3, P_3), $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, S_3), $(S_3, \overline{P}_3)$.

(b) P_4, $2K_2$, $P_3 + S_1$ and an edge.

A tree T can be partitioned into 2 stable sets, therefore at most one family contains all the stable sets, without loss of generality, this family is $\mathcal{F}_1$. If $\mathcal{F}_1$ is the set of all stable sets, then $\mathcal{F}_2$ is S_3 and $\overline{P_3}$-free, and therefore is the family $\mathcal{G}_1$. Otherwise both $\mathcal{F}_1$ and $\mathcal{F}_2$ contain edges. Assume without loss of generality, that $\mathcal{F}_2$ contains either P_3 or $\overline{P_3}$, then $\mathcal{F}_1$ is the family of all cliques and $\mathcal{F}_2$ is H-free for each $H \in \{P_4, 2K_2, P_3 + S_1\}$, and therefore by Claim 5.2.10, $\mathcal{F}_1$ is the family of all graphs which are a join of graphs which are disjoint union of a clique and a stable set. □

Theorem (1.1.16)**.** *For a tree $T \neq P_6$ such that $T \in \mathcal{T}^{\mathrm{npl}} \backslash (\mathcal{S} \cup \{P_6\})$, and let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be the $\mathcal{P}(T)$-free sequence, then we have the following case,*

(i) The families can be reindexed such that $\mathcal{F}_1$ is $\mathcal{G}_5$, and the rest of the families are the sets of all cliques.

Proof. By Lemmas 5.2.5 and 5.2.2, each tree $T \in \mathcal{T}^{\mathrm{npl}}$ can be partitioned into P_6 and $\mathrm{wpn}(T) - 2$ edges or in the case it is not a path it can be partitioned into P_3^* and $\mathrm{wpn}(T) - 2$

edges. Hence by Observations 5.2.8 and 5.2.7 and 5.2.9 and also Claim 5.2.6, each tree $T \in \mathcal{T}^{\mathrm{npl}} \setminus \mathcal{S}$ can be partitioned into,

(a) (P_3, P_3), $(P_3, \overline{P}_3)$, $(\overline{P}_3, \overline{P}_3)$, (S_3, S_3), $(S_3, \overline{P}_3)$, (S_3, P_3) and $\mathrm{wpn}(T) - 2$ edges.

(b) P_4, $2K_2$, $K_2 + S_2$, $P_3 + K_1$ and $\mathrm{wpn}(T) - 1$ edges.

(c) S_4 and $\mathrm{wpn}(T) - 1$ edges.

Assume that there is a family, without loss of generality, $\mathcal{F}_1$, which contains either of $\{S_3, P_3, \overline{P_3}\}$, then by (a) the rest of the families cannot contain either of the graphs from $\{S_3, P_3, \overline{P_3}\}$ and therefore are the sets of all cliques. Then by (b), (c) and Claim 5.2.10 $\mathcal{F}_1 = \mathcal{G}_4$.

If no family contains either of $\{S_3, P_3, \overline{P_3}\}$, then all the families are the sets of all cliques.

□

As in the case of trees in $\mathcal{T}^{\mathrm{pl}}$, we mention an observation and a few lemmas which are useful in the proofs of the exact structure of almost all T-free graphs for trees $T \in \mathcal{T}^{\mathrm{npl}}$ in Subsection 5.2.2

Observation 5.2.15. *Let $T \in \mathcal{T}^{\mathrm{npl}}$, then each $P(T)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ is properly arranged. Moreover, each of the families $\mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, is stable.*

The proofs of the following lemmas is similar to the corresponding proofs of Lemmas 5.2.16, 5.2.17 and 5.2.18.

Lemma 5.2.16. *Each of the graphs in $\{P_4, 2K_2, P_3 + K_1\}$ is $\frac{1}{4}$-universally extendable for all $T \in \mathcal{T}^{\mathrm{npl}}$. Moreover, $P_2 + S_2$ is $\frac{1}{4}$-universally extendable for all trees $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$ and S_4 is $\frac{1}{4}$-universally extendable for all trees $T \in \mathcal{T}^{\mathrm{npl}} \setminus \mathcal{S}$.*

Lemma 5.2.17. *The graph P_6 is $(2, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\mathrm{npl}}$.*

Lemma 5.2.18. *Either P_3^* is $(2, \frac{1}{4})$-universally extendable or P_8 is $(3, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$.*

Before we finish this subsection, we mention some additional structural properties of trees with a perfect matching and graphs in families $\mathcal{G}_1$ and $\mathcal{G}_4$.

Lemma 5.2.19. *Let $k \in \mathbb{N}$, $n \in \mathbb{N}$, $\tau \in \left(0, \frac{1}{2}\right)$ and $\gamma \in \left(0, \frac{\tau}{2^4}\right)$. Let $G \in (\mathcal{G}_1)_{\frac{n}{k}}$ such that $\overline{G}$ contains at most γn singleton components, and let $w \notin V(G)$ such that $|N(w) \cap V(G)| \geq \tau n$ and $|\overline{N}(w) \cap V(G)| \geq \tau n$. Then there is a collection $\mathcal{P}$ of disjoint copies of P_3 such that $|\mathcal{P}| \geq \frac{\tau n - \gamma n}{4}$ and one of the following cases happens,*

(i) w is adjacent to exactly one end (and not other vertices) of each copy of P_3 in $\mathcal{P}$, or

(ii) it is possible to choose $\mathcal{P}$ such that w is either not adjacent to any of the vertices of each copy of P_3 in $\mathcal{P}$ or w is adjacent to the centre vertex (an not other vertices) of each copy of P_3 in $\mathcal{P}$.

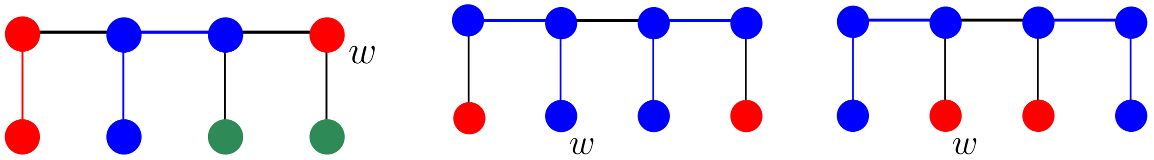
Proof. Let $G' = G[V(G) \setminus L]$ where L it the collection of singleton vertices in $\overline{G}$, let $\mathcal{A}$ be the collection of non-edges in G', let $A = \cup_{S \in \mathcal{A}} S = V(G')$, and let $N_1 = \overline{N}(w) \cap V(G')$ and $N_2 = N(w) \cap V(G')$. Then $|N_1| \geq \tau n - \gamma n$ and $|N_2| \geq \tau n - \gamma n$.

(i') There is a set of non-edges $\mathcal{A}' \subseteq \mathcal{A}$ such that $|\mathcal{A}'| = \frac{|N_1|}{2}$ and for each edge $\{a_1, a_2\} \in \mathcal{A}'$, $a_1 \in N_1$ and $a_2 \in N_2$.

(ii') There is a set of non-edges $\mathcal{A}' \subseteq \mathcal{A}$ such that $|\mathcal{A}'| = \frac{|N_1|}{2}$ and for each edge $\{a_1, a_2\} \in \mathcal{A}'$, $\{a_1, a_2\} \subseteq N_1$.

If case (i') happens then we attain the set $\mathcal{P}$ as in case (i) by taking all the copies of P_3 which are obtained from a non-edge in $\mathcal{A}'$ and a vertex in N_1 which is not a part of a non-edge in $\mathcal{A}'$. If case (ii') happens then we attain the set $\mathcal{P}$ as in case (ii) by taking all the copies of P_3 which are obtained from a non-edge in $\mathcal{A}'$ and either a vertex in N_1 or a vertex in N_2 which is not a part of a non-edge in $\mathcal{A}'$. □

Observation 5.2.20. *It is possible to choose a vertex $w \in V(P_4^*)$ such that either of the following cases holds.*



(a) in Observation 5.2.20 (b) in Observation 5.2.20 (c) in Observation 5.2.20

Figure 5.7: Partitions of P_4^* considered in Observation 5.2.20.

(a) There is partition of P_4^* into $S_2 + P_3 + \overline{P_3}$ such that w is a part of $\overline{P_3}$, and it is adjacent to one end of S_2 and one end of the P_3.

(b) There is partition of P_4^* into $S_2 + P_3 + P_3$ such that w is part of P_3, and it is not adjacent to S_2, and not adjacent to P_3.

(c) There is partition of P_4^* into $S_2 + P_3 + P_3$ such that w is a part of S_2, and it is not adjacent P_3 and adjacent to one of the ends of the P_3.

See in Figure 5.7 the partition as in Observation 5.2.20. Similarly to the proof of Lemma 5.2.19, it is possible to show the following.

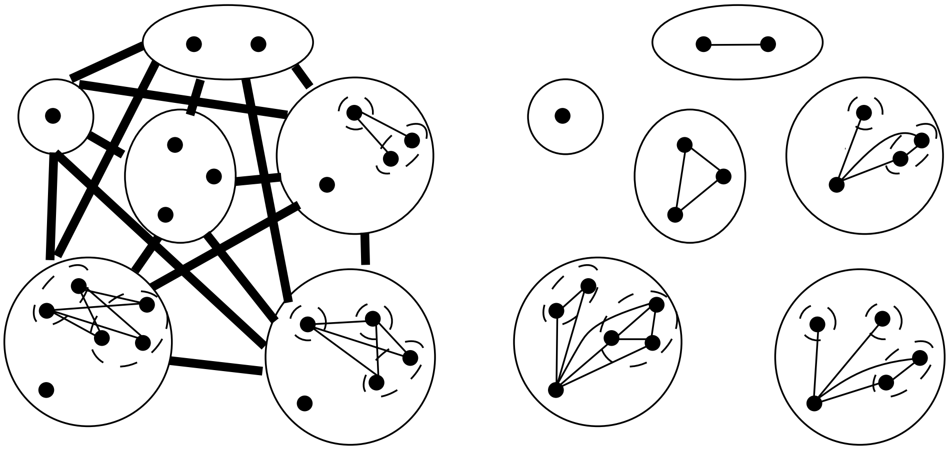
Lemma 5.2.21. *Let $k \in \mathbb{N}$ and $n \in \mathbb{N}$ large enough, let $\tau \in \left(0, \frac{1}{2}\right)$ and $\gamma \in \left(0, \frac{\tau}{2^4}\right)$. Let $G \in (\mathcal{G}_4)_{\frac{n}{k}}$ such that $\overline{G}$ contains at most γn vertices in components of size at least $(\log n)^2$, and let $w \notin V(G)$ such that $|N(w) \cap V(G)| \geq \tau n$ and $|\overline{N}(w) \cap V(G)| \geq \tau n$. Then there is a collection $\mathcal{P}$ of disjoint copies of P_3 such that $|\mathcal{P}| \geq \frac{\tau n - \gamma n}{(\log n)^2}$ and one of the following cases happens,*

(i) w is adjacent to exactly one end (and not other vertices) of each copy of P_3 in $\mathcal{P}$, or

(ii) w is not adjacent to any of the vertices of each copy of P_3 in $\mathcal{P}$.

5.2.2 Typical structure of T-free graphs for T with a perfect matching

In this subsection we finish the proof of Theorems 1.1.9, 1.1.11, 1.1.13, 1.1.15, and 1.1.17. We give separate proofs for trees in $\mathcal{T}^{\mathrm{pl}}$ and for trees in $\mathcal{T}^{\mathrm{npl}}$.



(a) $G \in \mathcal{G}_2$ (b) The complement $\overline{G}$

Figure 5.8: A graph G in $\mathcal{G}_2$ and its complement.

Trees in $\mathcal{T}^{\mathrm{pl}}$

We showed in Theorems 1.1.8 and 1.1.10 that for any $T \in \mathcal{T}^{\mathrm{pl}}$, the families $\mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$, in a $P(T)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ are either of the families $\mathcal{G}_\iota$ for $\iota \in [3]$ or the set of all cliques. We recall the definition of the families $\mathcal{G}_\iota$, $\iota \in [3]$.

The family $\mathcal{G}_1 = \mathrm{Forb}(\{S_3, \overline{P_3}\})$ is the family of graphs which are complete multi-partite graphs with parts of size at most 2. The complement of any graph in $\mathcal{G}_1$ is a disjoint union of a matching with a stable set.

The family $\mathcal{G}_2 = \mathrm{Forb}(\{P_4, 2K_2, K_2 + S_2\})$ is the family of graphs which are joins of graphs which are complete multi-partite graph and an isolated vertex. The complement of any graph in $\mathcal{G}_2$ is a disjoint union of clique-stars. Recall that a clique-star is a graph which is a complement of a disjoint union of a vertex and a complete multi-partite graph on at least two vertices. In particular, any clique-star contains either S_3 or $\overline{P_3}$. See in Figure 5.8 an example for a graph in $\mathcal{G}_2$ and its complement. In the figure, a thick edge between two sets of vertices represents the existence of all the possible edges between those two sets.

The family $\mathcal{G}_3 = \mathrm{Forb}(\{P_4, 2K_2, K_2 + S_2, S_4\})$ is the family of graphs which are joins of graphs which are complete multi-partite graph with parts of size at most 2 and an isolated vertex. Note that $\mathcal{G}_3 \subseteq \mathcal{G}_2$, and in particular $|(\mathcal{G}_3)_n| \leq |(\mathcal{G}_2)_n|$ for all $n \in \mathbb{N}$. Next we give bounds on the number of graphs in the families $\mathcal{G}_\iota$, for $\iota \in [3]$.

Theorem 5.2.22 ([32]). *Let $n \in \mathbb{N}$,*

$$|(\mathcal{G}_1)_n| = \left(1 + \mathrm{O}(n^{-1/2})\right) \left(\frac{n}{e}\right)^{n/2} \frac{e^{\sqrt{n}}}{(4e)^{1/4}}.$$

From the above we can derive the following bounds.

Corollary 5.2.23. *Let $n \in \mathbb{N}$,*

$$\left(\frac{n}{e}\right)^{n/2} \leq |(\mathcal{G}_1)_n| \leq n^{\frac{n}{2}}.$$

Let $\boldsymbol{(\mathcal{G}_1)^s_n} \subset (\mathcal{G}_1)_n$ be the set of graphs G on n vertices in $\mathcal{G}_1$ such that $\overline{G}$ contains at least s singleton components.

Lemma 5.2.24. *Let $n \in \mathbb{N}$ and $s \leq n$, then,*

$$(\mathcal{G}_1)^s_n \leq 2^n \cdot (n-s)^{\frac{n-s}{2}}.$$

Proof. There are at most 2^n ways to choose the singleton vertices. The graph induced on the rest $n-s$ vertices is in $\mathcal{G}_1$, therefore by Corollary 5.2.23 there are at most $(n-s)^{\frac{n-s}{2}}$ ways to choose such graphs. This gives the required bound. □

Let $\overline{G}_\iota = \{\overline{G} \mid G \in \mathcal{G}_\iota\}$, $\iota \in [3]$.

Lemma 5.2.25. *Let $n \in \mathbb{N}$,*

$$|\,(\mathcal{G}_2)_n\,| \leq 40^n \cdot \left(\frac{n}{\log\log n}\right)^n.$$

Proof. To make the presentation easier, we count the graphs in $(\overline{\mathcal{G}_2})_n$. Let $X \cup Y$ be a partition of $[n]$, and assume that the components induced on X are of size at most $\log n$ and the components induced on Y are of size at least $\log n$. There are 2^n ways to partition $[n]$ into the two sets X, Y.

Firstly we count the number of possible graphs on X. Let $|X| = x$, if $x \leq \sqrt[3]{n}$, then we bound the number of graphs on X by $2^x \cdot x^{2x} \leq 2^{\sqrt[3]{n}} n^{2/3\sqrt[3]{n}}$. Indeed, there are at most 2^x ways to choose the centres of the clique-stars on X, there are at most x^x ways to partition the vertices in X into the clique-stars and there are at most x^x ways to partition the vertices chosen for every clique-star into the different cliques of the clique-star. If $x \geq \sqrt[3]{n}$, then we count the number of possible graphs on X as following. Using the bound on the Bell numbers 1.3.2, there are at most $\left(\frac{x}{\log x}\right)^x \leq \left(\frac{n}{\frac{1}{3}\log n}\right)^x$ ways to partition X into the different components. Let C be a component on X, then again by 1.3.2 there are at most $\left(\frac{|C|}{\log(|C|)}\right)^{|C|} \leq \left(\frac{\log n}{\log\log n}\right)^{|C|}$ ways to partition it. Hence the number of ways to obtain a graph on X is at most

$$\left(\frac{n}{\frac{1}{3}\log n}\right)^x \cdot \prod_{C, V(C)\subseteq X} \left(\frac{\log n}{\log\log n}\right)^{|C|} = \left(\frac{n}{\frac{1}{3}\log n}\right)^x \cdot \left(\frac{\log n}{\log\log n}\right)^x = \left(\frac{3n}{\log\log n}\right)^x .$$

Secondly we count the number of possible graphs on Y. Let $|Y| = y$, let S_j be the union of the vertices in components of size between $2^j \log n$ and $2^{j+1}\log n$, and let $|S_j| = s_j$. There are at most $\frac{y!}{\Pi_{j=1}^{\log n - \log\log n} s_j!}$ ways to partition the vertices in Y into the S_j. Note that there are at least $\frac{s_j}{2^j \log n}$ components of sizes between $2^j\log n$ and $2^{j+1}\log n$. Therefore there are at most $\left(\frac{s_j}{2^j\log n}\right)^{s_j}$ ways to partition S_j into the components. Using again Theorem 1.3.2 and the fact that there are at most 2^{s_j} ways to choose the centres of the cliques stars, there are at most $2^{s_j}\left(\frac{2^{j+1}\log n}{\log(2^{j+1}\log n)}\right)^{s_j}$ ways to partition each of the components on S_j. Therefore for each S_j, we have at most $\left(\frac{s_j}{2^j\log n}\cdot\frac{2^{j+2}\log n}{\log(2^{j+1}\log n)}\right)^{s_j} = \left(\frac{4s_j}{\log(2^{j+1}\log n)}\right)^{s_j}$ ways to obtain a collection of clique-stars on S_j. Hence the number of ways to obtain a graph on Y is at most

$$\frac{y!}{\prod_{j=1}^{\log n - \log\log n} s_j!} \cdot \prod_{j=1}^{\log n - \log\log n} \left(\frac{4s_j}{\log(2^{j+1}\log n)}\right)^{s_j} \leq \left(\frac{4n}{\log\log n}\right)^y .$$

By taking the product over all x, y such that $x + y = n$, we get the required bound. □

Let $G \in \overline{\mathcal{G}_2}$, a matching $M \subset E(G)$ is **mild** if it is a matching in the subgraph G' of G which is obtained after removing the centres of all clique-stars which are not cliques from G. Let $\boldsymbol{(\mathcal{G}_2)_n^{c,n',n'',\ell}} \subset (\mathcal{G}_2)_n$ be the set of graphs G such that $\overline{G}$ has the following properties.

- $\overline{G}$ contains at most c clique-stars,

- the number of vertices in the clique-stars in $\overline{G}$ is at most n',

- there is a clique-star with at least n'' vertices in $\overline{G}$,
- there are at most ℓ edges in a mild matching in $\overline{G}$.

Similarly to before we denote $\boldsymbol{(\overline{\mathcal{G}_2})_n^{c,n',n'',\ell}} = \{\overline{G} \mid G \in (\mathcal{G}_2)_n^{c,n',n'',\ell}\}$. If there is no restriction on any of the values in the upper script, then we write a $*$ in the corresponding entry. In case that a clique-star is a clique we set any of its vertices to be the centre arbitrary.

Lemma 5.2.26. *Let $c, \ell, n', n'', n \in \mathbb{N}$, then*

$$|(\mathcal{G}_2)_n^{c,*,*,\ell}| \leq n^2 \cdot 2^{2n} \cdot (c+1)^n \cdot (3\ell)^{3\ell}. \tag{5.1}$$

$$|(\mathcal{G}_2)_n^{c,n',*,\ell}| \leq n^2 \cdot 2^{3n} \cdot (c+1)^{n'} \cdot (3\ell)^{3\ell}. \tag{5.2}$$

$$|(\mathcal{G}_2)_n^{*,*,n'',\ell}| \leq n \cdot 2^{3n} \cdot (n-n'')^{n-n''} \cdot (3\ell)^{3\ell}. \tag{5.3}$$

Proof. We consider the graphs in $(\overline{\mathcal{G}_2})_n^{c,*,*,\ell}$. For each $c' \leq c$ and $\ell' \leq \ell$, we count the number of graphs in with exactly c' clique-stars and exactly ℓ' edges in the mild matching. We have at most 2^n ways to choose the centres of the clique-stars and at most 2^n ways to choose the vertices for the mild matching.

Note that once we remove the centres of each of the clique-star we are left with a disjoint union of cliques, let G' be the resulting subgraph. The number of vertices in components of size at least 2 is at most $3\ell'$. Indeed, let $\mathcal{K}$ be the collection of all components of size at least 2. The number of such components is at most ℓ'. Assume that we remove a vertex from every odd component in $\mathcal{K}$, and let $\mathcal{K}'$ be the new collection of components of size at least 2. Each component $K \in \mathcal{K}'$ contributes exactly $\frac{|V(K)|}{2}$ edges to the mild matching. Therefore the number of vertices in the components in $\mathcal{K}'$ is $2\ell'$. To obtain $\mathcal{K}'$, we removed at most ℓ' vertices. Hence the total number of vertices in $\mathcal{K}$ is at most $3\ell'$. By the bound on the Bell numbers 1.3.2, there are at most $(3\ell')^{3\ell'}$ ways to partition $3\ell'$ vertices into any number of components of any size. There are at most 2^n ways to choose the singleton vertices in G'.

There are at most $(c'+1)^n$ ways to assign the vertices in G' to the different centres of the clique-stars together with the option that some of the components in G' are not assigned to any clique-star. Therefore in total there are at most $2^{3n} \cdot (c'+1)^n \cdot (3\ell')^{3\ell'}$ ways to obtain the required graph. Note that the bound increases with c' and ℓ'. Therefore summing over at most n^2 options for c' and ℓ', we get that the number of required graphs in $(\overline{\mathcal{G}_2})_n^{c,*,*,\ell}$ is at most

$$n^2 \cdot 2^{2n} \cdot (c+1)^n \cdot (3\ell)^{3\ell}.$$

Now we show the bound on $(\mathcal{G}_2)_n^{c,n',*,\ell}$. The number of ways to choose the n' vertices which are in the clique-stars is at most 2^n. As before there are at most 3ℓ vertices in G'. We have at most $(3\ell)^{3\ell}$ way to partition those vertices into the different components. We have at most $(c+1)^{n'}$ ways to assign the n' vertices to the different centres together with the option of not assigning those vertices to any of the centres. Therefore the number of graphs in $(\mathcal{G}_2)_n^{c,n',*,\ell}$ is at most

$$n^2 \cdot 2^{3n} \cdot (c+1)^{n'} \cdot (3\ell)^{3\ell}.$$

Now we show the bound on $(\mathcal{G}_2)_n^{*,*,n'',\ell}$. As before there are at most 3ℓ vertices in G'. There are at most $(3\ell)^{3\ell}$ way to partition those vertices into the different components. Let K be the clique-star which contains n'' vertices. We have at most 2^n to choose those vertices. On the remaining $n - n''$ vertices we can have at most $n - n''$ clique-stars. The number of ways to partition those remaining vertices into the clique-stars is at most $(n-n'')^{n-n''}$. Therefore the number of graphs in $(\mathcal{G}_2)_n^{*,*,n'',\ell}$ is at most

$$n \cdot 2^{3n} \cdot (n-n'')^{n-n''} \cdot (3\ell)^{3\ell}.$$

□

Let $\boldsymbol{(\mathcal{G}_3)_n^c} \subset (\mathcal{G}_3)_n$ be the family of graphs G on $[n]$ such that $\overline{G}$ contains at most c clique-stars.

Lemma 5.2.27. *Let $n \in \mathbb{N}$ and $c \in [n]$, then*

$$|(\mathcal{G}_3)_n^c| \leq 2^n \cdot (c+1)^n \cdot n^{\frac{n}{2}}.$$

Proof. There are at most 2^n ways to choose the c centres of the clique-stars. Let $G \in (\mathcal{G}_3)_n^c$ and let G' be the resulting graph after removing the centres of the clique-stars from $\overline{G}$. Then

$\overline{G}'$ is a disjoint union of edges and vertices. By Theorem 5.2.22, there are at most $n^{\frac{n}{2}}$ ways to partition n vertices into edges and vertices. There are at most $(c+1)^n$ ways to assign the vertices of G' into the different clique-stars together with the option that the vertices are not assigned to any clique-star. By taking the product of all of the above, we get the required bound. □

Observation 5.2.28. *Let $T \in \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$ and Π a partition of $[n]$, then*

$$|F(T,\Pi)| \geq |(\mathcal{G}_1)_n| \geq \left(\frac{n}{e}\right)^{n/2}.$$

Observation 5.2.29. *Let $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$ and let Π be a $\rho/4$-almost equal partition for some $\rho > 0$. Let $n_1 := |\pi_1|$ and $n_2 := |\pi_2|$, then*

$$|F(T,\Pi)| = |(\mathcal{G}_1)_{n_1}| \cdot |(\mathcal{G}_1)_{n_2}| \geq \left(\frac{n_1}{e}\right)^{\frac{n_1}{2}} \cdot \left(\frac{n_2}{e}\right)^{\frac{n_2}{2}} \geq \left(\frac{n/\operatorname{wpn}(T) - n^{\rho/4}}{e}\right)^{(n/\operatorname{wpn}(T) - n^{\rho/4})}.$$

Observation 5.2.30. *All the following pairs of graphs can be extended to P_3^* with a proper choice of edges between the graphs in the pair,*

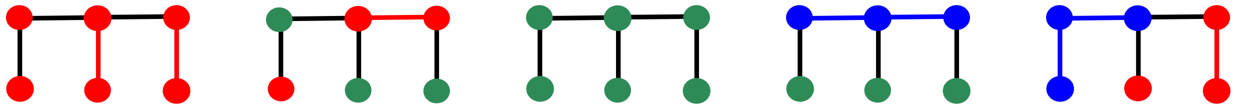
(a) $(\overline{P_3}, \overline{P_3})$ with an edge between their centres, see Figure 5.9a.

(b) $(\overline{P_3}, S_3)$ with either an edge between the centre of the $\overline{P_3}$ and one of vertices of S_3 or without such edge, see Figure 5.9b.

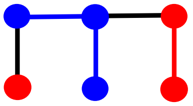
(c) (S_3, S_3) with an edge fixed between some vertex from one of the S_3 to some vertex from the other S_3 or without such edge, see Figure 5.9c.

(d) (P_3, S_3) with either an edge between the centre of the P_3 and one of vertices of S_3 or with an edge between one of the ends of the P_3 and one of vertices of S_3, see Figure 5.9d.

(e) $(P_3, \overline{P_3})$ with either an edge between the centre of the $\overline{P_3}$ and one of the ends of P_3 or without such edge, see Figure 5.9e. Moreover, $(P_3, \overline{P_3})$ with an edge between the centre of P_3 and one of the ends of the edge of the $\overline{P_3}$, see Figure 5.9f.



(a) $\overline{P_3}$ and $\overline{P_3}$ (b) $\overline{P_3}$ and S_3 (c) S_3 and S_3 (d) P_3 and S_3 (e) P_3 and $\overline{P_3}$



(f) P_3 and S_3

Figure 5.9: Partitions of P_3^* considered in Observation 5.2.30.

Lemma 5.2.31. *Let K_1 and K_2 be such that $\overline{K_1}$ and $\overline{K_2}$ are the disjoint union of a clique-star S_1 (resp. S_2) on k_1 (resp. k_2) vertices together with k_1 (resp. k_2) vertices (there is no restriction on the graph on those vertices). The number of ways to choose edges between $V(K_1)$ and $V(K_2)$ without creating an induced copy of P_3^* is at most*

$$2^{2k_1\cdot 2k_2}\cdot 2^{-\min\left\{\frac{k_1-1}{2^6},\frac{k_2-1}{2^6}\right\}}.$$

Proof. Let c_1 (resp. c_2) be the centre of the clique-star S_1 (resp S_2). Note that $\overline{S_1}$ contains at least $\frac{k_1-1}{2}$ disjoint edges or disjoint non-edges. Similarly, $\overline{S_2}$ contains at least $\frac{k_2-1}{2}$ disjoint edges or disjoint non-edges. We have three possible cases: (1) $\overline{S_1}$ contains at least $\frac{k_1-1}{2}$ disjoint edges and $\overline{S_1}$ contains at least $\frac{k_2-1}{2}$ disjoint edges; (2) $\overline{S_1}$ contains at least $\frac{k_1-1}{2}$ disjoint non-edges and $\overline{S_2}$ contains at least $\frac{k_2-1}{2}$ disjoint non-edges; (3) without loss of generality, $\overline{S_1}$ contains at least $\frac{k_1-1}{2}$ disjoint edges and $\overline{S_1}$ contains at least $\frac{k_2-1}{2}$ disjoint non-edges.

First assume that there is an edge between c_1 and c_2. If we are in case (1), we use part (a) in Observation 5.2.30. The graphs K_1 and K_2 are fixed and we are choosing edges between $V(K_1)$ and $V(K_2)$ with probability $\frac{1}{2}$. Let E_1 be the set of disjoint edges in K_1 such that c_2 is not adjacent to any end of each such edge, and let E_2 be the set of disjoint edges in K_2 such that c_1 is adjacent to exactly one end of each such edge. The expected size of E_1 is $\frac{k_1-1}{2^2\cdot 2}$, and the expected size of E_2 is $\frac{k_2-1}{2\cdot 2}$. By Chernoff bound 1.3.6, the probability that either $|E_1| \le \frac{k_1-1}{2^3\cdot 2}$ or $|E_2| \le \frac{k_2-1}{2^2\cdot 2}$ is $e^{-\frac{k_1-1}{2^5}}+e^{-\frac{k_2-1}{2^4}}$. The number of graphs where $|E_1| \ge \frac{k_1-1}{2^3\cdot 2}$ and

$|E_2| \geq \frac{k_1-1}{2^2\cdot 2}$ is at most $2^{2k_1\cdot 2k_2}\cdot\left(\frac{2^4-1}{2^4}\right)^{\frac{k_1-1}{2^3\cdot 2}\cdot\frac{k_1-1}{2^2\cdot 2}}$, this is because we must forbid at least one edge arrangement between every $e_1 \in E_1$ and $e_2 \in E_2$, otherwise we get a copy of P_3^*. Hence in total, the number of possible graphs is at most $2^{2k_1\cdot 2k_2}\cdot 2^{-\min\left\{\frac{k_1-1}{2^6},\frac{k_2-1}{2^6}\right\}}$.

In all the other cases the argument is very similar. If we are in case (2), then we use part (c) of Observation 5.2.30. If we are case (3), then we use part (b) of Observation 5.2.30.

If there is no edge between c_1 and c_2, if we are in case (1), then we use the first arrangement of part (e) of Observation 5.2.30. We can find many disjoint P_3 either in K_1 or K_2 by using the vertices which are not part of the clique-stars. If we are in case (2), then as before we use part (c) of Observation 5.2.30. If we are case (3), then we use part (b) of Observation 5.2.30. □

We always apply the above lemma to two collections of clique-stars.

Lemma 5.2.32. *Let $G_1, G_2 \in \mathcal{G}_2$ be two disjoint graphs. Let $|V(G_1)| = n_1$, $|V(G_2)| = n_2$ and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be maximum collections of disjoint clique-stars in $\overline{G}_1$ and $\overline{G}_2$, respectively. Let $s_1 = |\mathcal{S}_1|, s_2 = |\mathcal{S}_2|$ and $S_1 = \cup_{S\in\mathcal{S}_1} S, S_2 = \cup_{S\in\mathcal{S}_2} S$. Let $n_1' \leq \frac{n_1}{4}$. We make the following assumptions.*

(i) The largest clique-star in $\mathcal{S}_1$ contains at most $n_1 - n_1'$ vertices,

(ii) $n_1' \leq |S_1| \leq \frac{n_1}{4}$, and

(iii) every clique-star in $\mathcal{S}_2$ is larger than any clique-star in $\mathcal{S}_1$.

The number of ways to choose edges between $V(G_1)$ and $V(G_2)$ without creating an induced copy of P_3^* is at most $2^{n_1\cdot n_2}\cdot 2^{-cn_1'\cdot s_2}$ for some constant $c > 0$.

Proof. We define two collections $\mathcal{K}_1$ and $\mathcal{K}_2$ of subgraphs of G_1 and G_2, respectively, so we can apply Lemma 5.2.31 to each pair of sets (K_1, K_2) such that $K_1 \in \mathcal{K}_1$ and $K_2 \in \mathcal{K}_2$.

First we define a set $\mathcal{K}_1$ of subgraphs of G_1. Let L be the clique-star with most vertices in $\mathcal{S}_1$. If $|V(L)| \geq n_1'$, then we define a component K to be a subgraph of L which is a clique-star with n'_1 vertices together with a set of some n'_1 vertices in $V(G) \setminus V(L)$. Note that $|V(G) \setminus V(L)| \geq n'_1$ due to assumption (i). In this case $\mathcal{K}_1 = \{K\}$.

Next assume that $|V(L)| < n'_1$. In this case, we can find a collection of clique-stars which contains between n'_1 and $2n'_1$ vertices. Indeed, let $C_1, C_2, ..., C_k = L$ be the ordering of the clique-stars in $\mathcal{S}_1$ in increasing order by the number of vertices. Let $t \in [k]$ be such that $\sum_{i=1}^{t} |V(C_i)| < n'_1$ and $\sum_{i=1}^{t+1} |V(C_i)| \geq n'_1$. By assumption (ii), $|S_1| = \sum_{i=1}^{k} |V(C_i)| \geq n'_1$, therefore $t < k$. Moreover, $|V(C_t)| \leq |V(L)| \leq n'_1$. Therefore $n'_1 \leq \sum_{i=1}^{t+1} |V(C_i)| \leq 2n'_1$. Let V' be a set of vertices in $V(G) \setminus \cup_{i=1}^{t+1} V(C_i)$ such that $|V'| = \sum_{i=1}^{t+1} |V(C_i)|$, such a set exists because $|S_1| \leq \frac{n_1}{4}$. We partition V' into parts of sizes $|V(C_1)|, |V(C_2)|, ..., |V(C_{t+1})|$, let $V'_1, V'_2, ..., V'_{t+1}$ be the resulting partition where $|V'_i| = |V(C_i)|$ for $i \in [t+1]$. We define the set K_i to be $G[V(C_i) \cup V'_i]$. Let $\mathcal{K}_1$ be the collection of sets K_i, $i \in [t+1]$.

Next we define the set $\mathcal{K}_2$. Let $C_1, C_2, ..., C_k$ be the ordering of the clique-stars in $\mathcal{S}_2$ in increasing order by the number of vertices. We define a set K_{2i}, $1 \leq i \leq k/2$, by taking C_{i-1} together with $|V(C_{i-1})|$ vertices from C_i. We can do that because $|V(C_i)| \geq |V(C_{i-1})|$ for all $1 \leq i \leq k/2$.

We apply Lemma 5.2.31 to each pair of a set from $\mathcal{K}_1$ and a set from $\mathcal{K}_2$. This gives the required bound. □

We recall the structural properties of trees in $\mathcal{T}^{\text{pl}}$ which are needed in the proof of the exact structure of almost all T-free graphs.

(A) Each of the graphs in $\{P_4, K_2+S_2\}$ is $\frac{1}{4}$-universally extendable for all $T \in \mathcal{T}^{\text{pl}}$. Moreover, each of the graphs in $\{2K_2, S_4\}$ is $\frac{1}{4}$-universally extendable for all trees $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$.

(B) P_3^* is $(2, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\text{pl}} \setminus \{P_3^*\}$.

(C) P_4^* is $(3, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$.

Those properties are shown in Lemmas 5.2.12, 5.2.13 and 5.2.14. Using those properties we give a few important corollaries to the general claims which we gave in Section 3.3.

Let $K \in \mathbb{N}$ and $\varepsilon > 0$ be the constants from Theorem 3.2.1 applied with $\mathrm{Forb}(T)$ and $\delta > 0$ sufficiently small. Let $\varepsilon' > 0$ be the constant from Theorem 3.2.3 applied with K. Let $\lambda \in [0, \frac{\min\{\varepsilon, \varepsilon'\}}{8})$.

Let $n \in \mathbb{N}$ be large enough and let $\rho > 0$ be the constant which we get from Theorem 3.2.2 for this n and $\xi > 0$ such that $2\xi \log \frac{1}{\xi} < \log e \cdot \frac{\lambda}{2^{3 \cdot 2^{t^2+3} \cdot t \cdot \mathrm{wpn}(T)}}$ where $t = V(T)$. Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(T)})$ be a $\rho/4$-almost equal partition of $[n]$. We fix the partition Π for all of the following discussion. Let $\boldsymbol{n_i}= |\pi_i|$, $i \in [\mathrm{wpn}(T)]$.

Let G be a Π-conformal graph. Let $\mathcal{Y}(\Pi, G, i) = \mathcal{Y}(\Pi, G, \frac{1}{4}, i, \frac{1}{2t^2+1})$ be the collection of sets obtained by adding greedily sets S which are $(\Pi, \frac{1}{4}, i, \frac{1}{2t^2+1})$-linearly extremal. Let $\mathcal{Y}(G) = \cup_{i=1}^{\mathrm{wpn}(T)} \mathcal{Y}(\Pi, G, i)$, $Y = Y(G) = \cup_{Y \in \mathcal{Y}(G)} Y$ and $y = y(G) = |Y(G)|$ as defined in Section 3.3.

Corollary 5.2.33 (to Lemma 3.3.9)**.** *Let $T \in \mathcal{T}^{\mathrm{pl}}$, the number of Π-conformal graphs G, such that for some $i \in [\mathrm{wpn}(T)]$, $G[\pi_i \setminus Y(G)]$ contains an induced graph isomorphic to H' where $H' \in \{P_4, K_2 + S_2\}$ is much smaller than the number of Π-good graphs.*

Moreover, if $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, then the number of Π-conformal graphs G such that additionally there is an index $i \in [\mathrm{wpn}(T)]$ where $G[\pi_i \setminus Y(G)]$ contains an induced graph isomorphic to H' where $H' \in \{2K_2, S_4\}$ is much smaller than the number of Π-good graphs.

Corollary 5.2.34 (to Corollary 3.3.10 and Theorem 2.2.4)**.** *Let $T \in \mathcal{T}^{\mathrm{pl}}$, there is a constant $C(T) \geq 0$ which depends only on T such that the number of Π-conformal graphs where $y \geq C(T) \log n$ is much smaller than the number of Π-good graphs.*

Corollary 5.2.35 (to Lemma 3.3.12)**.** *Let $T \in \mathcal{T}^{\mathrm{pl}}$, the number of Π-conformal graphs where there are indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, and there are subgraphs J_{i_1} and J_{i_2} in $G[\pi_{i_1} \setminus Y(G)]$ and $G[\pi_{i_2} \setminus Y(G)]$, respectively, such that $G[V(J_{i_1}) \cup V(J_{i_2})]$ is isomorphic to P_3^* is much smaller than the number of Π-good graphs.*

Corollary 5.2.36 (to Lemma 3.3.13)**.** *Let $T \in \mathcal{T}^{\mathrm{pl}}$, and let (J_1, J_2) be a partition of P_3^*. The number of Π-conformal graphs where there indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, such that $G[\pi_{i_1} \setminus Y(G)]$ contains at least $f_1(n)$ disjoint copies of a graph J_1 and $G[\pi_{i_2} \setminus G(Y)]$ contains at least $f_2(n)$*

disjoint copies of a graph J_2 *such that* $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$*, is much smaller than the number of* Π*-good graphs.*

Corollary 5.2.37 (to Lemma 3.3.13)**.** *Let* $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$*, and let* (J_1, J_2, J_3) *be a partition of* P_4^**. The number of* Π*-conformal graphs where there indices* i_1, i_2, i_3 *in* $[\mathrm{wpn}(T)]$*, such that* $G[\pi_{i_1} \setminus G(Y)]$ *and* $G[\pi_{i_2} \setminus G(Y)]$ *contain at least* $f_1(n)$ *disjoint copies of a graph* J_1 *and* J_2 *respectively, and* $G[\pi_{i_3} \setminus G(Y)]$ *contains at least* $f_2(n)$ *disjoint copies of a graph* J_3 *such that* $f_1(n) \geq f_2(n)$ *and* $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$*, is much smaller than the number of* Π*-good graphs.*

In every step in the following proof we define a property of Π-conformal graphs. We show that the number of graphs having this property is much smaller than the number of Π-good graphs. This allows us to focus on Π-bad graphs without this first property. For this subset we again define a new property and again show that the number of Π-conformal graphs with this new property is much smaller than the number of Π-good graphs. We continue in this manner until we conclude that a typical Π-bad graph does not have all the above properties and has a very specific structure. We finish by showing that also the number of Π-bad graphs in this final set is much smaller than the Π-good graphs.

Let $\mathscr{C}$ be the set of all Π-conformal graphs and let $\mathscr{G}$ be the set of all Π-good graphs. Let $\mathscr{C}' \subset \mathscr{C}$ be the set of all Π-conformal graphs G with the following properties.

(a) For all $T \in \mathcal{T}^{\mathrm{pl}}$, $G[\pi_i \setminus Y(G)] \in \mathrm{Forb}(\{P_4, K_2 + S_2\})$, $i \in [\mathrm{wpn}(T)]$, and for all $T \in \mathcal{T}^{\mathrm{pl}} \setminus \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$, $G[\pi_i \setminus Y(G)] \in \mathrm{Forb}(\{P_4, 2K_2, K_2 + S_2, S_4\}) = \mathcal{G}_3$, $i \in [\mathrm{wpn}(T)]$.

(b) $y(G) \leq C(T) \log n$ for some constant $C(T) \geq 0$,

(c) There are no indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, where there are subgraphs J_{i_1} and J_{i_2} in $G[\pi_{i_1} \setminus Y(G)]$ and $G[\pi_{i_2} \setminus Y(G)]$, respectively, such that $G[J_{i_1} \cup J_{i_2}]$ is isomorphic to P_3^*,

(d) There are no indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, such that $G[\pi_{i_1} \setminus Y(G)]$ contains at least $f_1(n)$ disjoint copies of a graph J_1 and $G[\pi_{i_2} \setminus Y(G)]$ contains at least $f_2(n)$ disjoint copies of a graph J_2 such that the following holds. The collection (J_1, J_2) is a partition of P_3^* and, $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$.

(e) For a tree $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$ there are no indices $i_1 \neq i_2 \neq i_3 \in [\text{wpn}(T)]$, such that $G[\pi_{i_1} \setminus Y(G)]$ and $G[\pi_{i_2} \setminus Y(G)]$ contain at least $f_1(n)$ disjoint copies of a graph J_1 and J_2 respectively, and $G[\pi_{i_3} \setminus Y(G)]$ contains at least $f_2(n)$ disjoint copies of a graph J_3 such that the following holds. The collection (J_1, J_2, J_3) is a partition of P_4^*, $f_1(n) \geq f_2(n)$ and $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \text{o}(1)$.

Lemma 5.2.38. *The number of graphs in $\mathscr{C} \setminus \mathscr{C}'$ is much smaller than the number of graphs in $\mathscr{G}$.*

Proof. This is a direct corollary to Corollaries 5.2.33, 5.2.34, 5.2.35, 5.2.36 and 5.2.37. $\square$

Let $\mathscr{F}$ be a family of graphs G which are defined with respect to the properties of the subgraphs $G[\pi_i]$, $i \in [\text{wpn}(T)]$. Recall that $\mathcal{F}(\mathscr{F}) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ is a collection of families where $\mathcal{F}_i = \cup_{G \in \mathscr{F}} G[\pi_i]$, $i \in [\text{wpn}(T)]$.

Before we give the proof for our main theorem of this section we give one more lemma. Let $\mathscr{C}'' \subseteq \mathscr{C}'$ be a collection of Π-conformal graphs G such that there is an $i \in [\text{wpn}(T)]$, without loss of generality $i = 1$, such that $G[\pi_1 \setminus Y(G)]$ contains a stable set of size at least $\frac{3n_1}{4}$. Note that by the structure of the graphs in a $P(T)$-free sequence, there can be at most one such index and also only for trees in $\mathcal{T}^{\text{pl}}_{\text{star}}$.

Lemma 5.2.39. *The number of graphs in $\mathscr{C}''$ is much smaller than the number of graphs in $\mathscr{G}$.*

Proof. Let $G \in \mathscr{C}''$, by the definition of the family, $G[\pi_1 \setminus Y(G)]$ contains a stable set of size at least $\frac{3n_1}{4} \geq \frac{3n}{2^3 \, \text{wpn}(T)}$. Hence we can find in $G[\pi_1 \setminus Y(G)]$ a set $\mathcal{J}_1$ of at least $\frac{n}{2^5 \, \text{wpn}(T)}$ disjoint copies of S_3. Hence by property (d) of graphs in $\mathscr{C}'$, for all $i \geq 2$, $G[\pi_i \setminus Y(G)]$ contains at most $(\log n)^2$ disjoint copies of either S_3, P_3 or $\overline{P_3}$. Let $X_i \subseteq (\pi_i \setminus Y(G))$ be the vertex set of the collection of those copies in $\pi_i \setminus Y(G)$. Note that $G[(\pi_i \setminus Y(G)) \setminus X_i]$ is a clique.

We bound the number of graphs with the above properties and with $y \leq C(T) \log n$. Let $Y \subset [n]$ be some set of size at most $C(T) \log n$ and let $\pi_i' = \pi_i \setminus Y$, $i \in [\text{wpn}(T)]$.

Let $\mathcal{B}_1 \subseteq \mathscr{C}''$ be the family of graphs where there is an index $i \geq 2$, such that there is a vertex $v \in X_i$ with $|\overline{N}(v) \cap (\pi_i' \setminus X_i)| \geq n^{\frac{1}{2}+\varepsilon}$ for some $\varepsilon > 0$. By our assumptions $\{v\} \notin \mathcal{Y}$, therefore by the definition of the set $\mathcal{Y}$, there is a set $\mathcal{J}_1' \subseteq \mathcal{J}_1$ of disjoint copies of S_3 such that v is adjacent to only one vertex of each $J \in \mathcal{J}_1'$ and $|\mathcal{J}_1'| \geq \frac{n}{2^{t^2+6}\,\mathrm{wpn}(T)}$. Let $\mathcal{E}_i$ be a maximal collection of disjoint edges in $\overline{N}(v) \cap (\pi_i' \setminus X_i)$. There is an edge arrangement that cannot appear between any S_3 in $\mathcal{J}_1'$ and an edge $e \in \mathcal{E}_i$, otherwise we get a copy of P_3^* contradicting property (c). Let $\mathcal{F}(\mathcal{B}_1) = (\mathcal{F}_1^1, \mathcal{F}_2^1, ..., \mathcal{F}_{\mathrm{wpn}(T)}^1)$, then by property (a) of graphs in $\mathscr{C}$ and Theorem 2.2.4 about the number of P_4-free graphs, for each $i \in [\mathrm{wpn}(T)]$, $|\mathcal{F}_i^1| \leq 2^{3n \log n}$. Therefore using Lemma 3.3.8, we can bound the number of graphs in $\mathcal{B}_1$ in this case by

$$2^{m(\Pi)} \cdot 2^{3n \log n} \cdot 2^{(b+3)n} \cdot 2^{-c(T)ny} \cdot \left(\frac{2^6-1}{2^6}\right)^{\frac{n}{2^{t^2+6}\,\mathrm{wpn}(T)} \cdot \frac{n^{\frac{1}{2}+\varepsilon}}{2}}$$

which is much smaller than the number of Π-good graphs. Therefore for each $i \geq 2$, for each $v \in X_i$, $|\overline{N}(v) \cap (\pi_i' \setminus X_i)| \leq n^{\frac{1}{2}+\varepsilon}$ for any $\varepsilon > 0$.

Let $\mathcal{F}(\mathscr{C}'' \setminus \mathcal{B}_1) = (\mathcal{F}_1^2, \mathcal{F}_2^2, ..., \mathcal{F}_{\mathrm{wpn}(T)}^2)$. We have that $|\mathcal{F}_1^2| \leq 2^n \cdot \left(\frac{n_1}{4}\right)^{\frac{n_1}{4}}$. Indeed, there are at most 2^n ways to choose the stable set of size at least $\frac{3n_1}{4} \geq \frac{3n}{2^3\,\mathrm{wpn}(T)}$ in $G \in \mathcal{F}_1^2$. By Lemma 5.2.25, the number of ways to obtain the graph on the vertex set which is not a part of the largest stable set is at most $\left(\frac{n_1}{4}\right)^{\frac{n_1}{4}}$.

Let $i \geq 2$, there are at most $\binom{n_i}{(\log n)^2}$ ways to choose the set X_i, by Lemma 5.2.25, there are at most $(\log n)^{2(\log n)^2} = 2^{2(\log n)^2 \cdot \log\log n}$ to choose the graph on X_i. By our assumptions, there are at most $\binom{n_i}{n^{\frac{1}{2}+\varepsilon}}^{|X_i|} \leq n_i^{n^{\frac{1}{2}+\varepsilon} \cdot |X_i|}$ ways to choose the edge arrangement between X_i and $\pi_i \setminus X_i$. Therefore for each $i \geq 2$, $|\mathcal{F}_i^2| \leq \binom{n_i}{(\log n)^2} \cdot 2^{2(\log n)^2 \cdot \log\log n} \cdot n_i^{n^{\frac{1}{2}+\varepsilon} \cdot |X_i|}$. Using Lemma 3.3.8 the number of graphs in $\mathscr{C}'' \setminus \mathcal{F}(\mathcal{B}_1)$ is at most

$$\begin{aligned} &2^{m(\Pi)} \cdot 2^{(b+3)n} \cdot 2^{-c(T)ny} \cdot 2^{n_1} \left(\frac{n_1}{4}\right)^{\frac{n_1}{4}} \cdot \prod_{i \geq 2} \binom{n_i}{(\log n)^2} \cdot 2^{2(\log n)^2 \cdot \log\log n} \cdot n_i^{n^{\frac{1}{2}+\varepsilon} \cdot |X_i|} \\ &\leq 2^{m(\Pi)} \cdot 2^{(b+3)n} \cdot 2^n \left(\frac{n_1}{4}\right)^{\frac{n_1}{4}} \cdot 2^{\mathrm{wpn}(T) n^{\frac{1}{2}+\varepsilon} (\log n)^3}, \end{aligned}$$

using that we choose $\varepsilon < \frac{1}{2}$ and the lower bound of $\left(\frac{n_1}{e}\right)^{\frac{n_1}{2}} \cdot \left(\frac{n_2}{e}\right)^{\frac{n_2}{2}}$ on the number of Π-good graphs from Observation 5.2.29, we conclude that the above is also much smaller than the number of Π-good graphs. $\square$

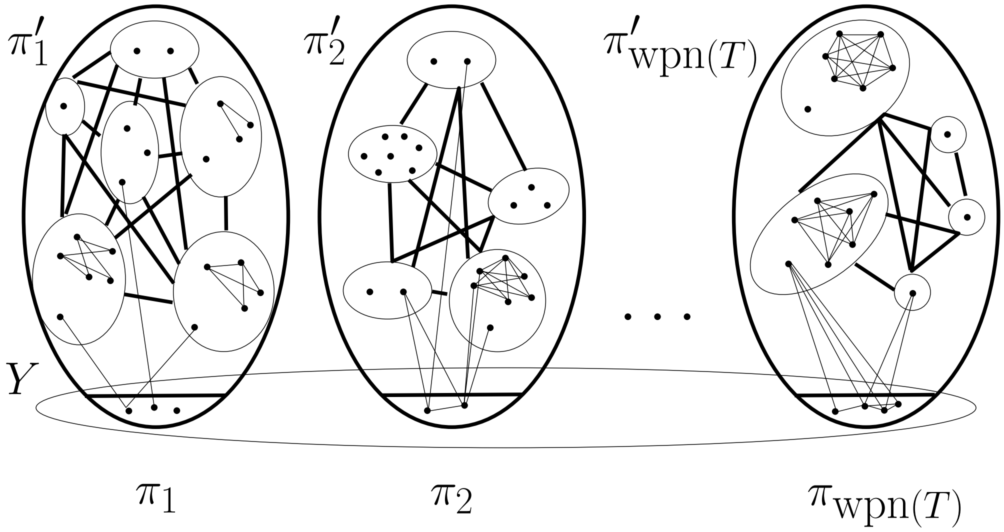


Figure 5.10: A sketch of a graph $G \in \mathscr{B}$. The edges between the parts are not drawn.

We focus on graphs in $\mathscr{C}' \setminus \mathscr{C}''$. By the definition of this family, for each $i \in [\mathrm{wpn}(T)]$, $G[\pi_i \setminus Y(G)]$ contains at least $\frac{n_1}{4}$ disjoint edges. Hence, similarly to the proof of Lemma 3.3.9, in almost all graphs in the family, for each $i \in [\mathrm{wpn}(T)]$, $G[\pi_i \setminus Y(G)]$ is also $2K_2$-free. Hence we can assume that $G[\pi_i \setminus Y(G)] \in \mathcal{G}_2$, $i \in [\mathrm{wpn}(T)]$, and $\overline{G}[\pi_i \setminus Y(G)]$ is a disjoint union of clique-stars with edges and singleton vertices. For each $i \in [\mathrm{wpn}(T)]$, we define $\boldsymbol{c_i = c_i(\overline{G}[\pi_i \setminus Y(G)])}$ to be the number of clique-stars in $\overline{G}[\pi_i \setminus Y(G)]$ and similarly we define $\boldsymbol{\ell_i = \ell_i(\overline{G}[\pi_i \setminus Y(G)])}$ to be the maximum number of edges in the mild matching in $\overline{G}[\pi_i \setminus Y(G)]$. Let $G \in \mathscr{C}' \setminus \mathscr{C}''$, we denote by $\boldsymbol{\pi_i'}= \pi_i \setminus Y(G)$, $i \in [\mathrm{wpn}(T)]$.

Let $\mathscr{B} \subseteq \mathscr{C}' \setminus \mathscr{C}''$ be the collection of all Π-bad graphs in $\mathscr{C}' \setminus \mathscr{C}''$. Note that we will use the fact that we are considering Π-bad graphs only towards the end of the proof. See Figure 5.10 for a sketch of an example of a graph in $\mathscr{B}$. In particular, in Figure 5.10, $c_1 = 4, \ell_1 = 6$, $c_2 = 3, \ell_2 = 8$, $c_{\mathrm{wpn}(T)} = 2, \ell_{\mathrm{wpn}(T)} = 2$.

Our first goal is to show that the number of graphs in $\mathscr{B}$ which do not have an $i \in [\mathrm{wpn}(T)]$, such that $\overline{G}[\pi_i \setminus Y(G)]$ contains a large mild matching is much smaller than the number in $\mathscr{G}$. Let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$ and $\boldsymbol{\mathscr{B}_1 = \mathscr{B}_1(\alpha)} \subset \mathscr{B}$ be the set of graphs in $\mathscr{B}$ such that for each $i \in [\mathrm{wpn}(T)]$, $\ell_i \leq \alpha n$.

Lemma 5.2.40. *Let $T \in \mathcal{T}^{\mathrm{pl}}$ and $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$. The number of graphs in $\mathscr{B}_1$ is much* *smaller than the number of* Π*-good graphs* $\mathscr{G}$*.*

Proof. We start by listing the different subsets of graphs in $\mathscr{B}_1$ that we consider during the proof. We show that the number of graphs in each of the following subsets is much smaller than the number of graphs in $\mathscr{G}$. We finish by showing the number of graphs in $\mathscr{B}_1$ which are not in any of the following families is also much smaller the number of graphs in $\mathscr{G}$.

Let $\beta \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^2}\right)$. Let $\mathcal{B}_1 = \mathcal{B}_1(\beta) \subset \mathscr{B}_1$ be the set of graphs G which have at most n^β clique-stars $\overline{G}[\pi'] \cup \overline{G}[\pi'_2] \cup ... \cup \overline{G}[\pi'_{\mathrm{wpn}(T)}]$.

Let $\mathcal{S}$ be the set of the n^β largest clique-stars in $\overline{G}[\pi'_1] \cup \overline{G}[\pi'_2] \cup ... \cup \overline{G}[\pi'_{\mathrm{wpn}(T)}]$. Let S be the union of the vertices in the components in $\mathcal{S}$. Let $\mathcal{B}_2 \subset \mathscr{B}_1 \backslash \mathcal{B}_1$ be the set of Π-bad graphs where each subgraph $\overline{G}[\pi'_i \setminus S]$, $i \in [\mathrm{wpn}(T)]$, has at least one of the following properties,

(1) it contains at most n^β clique-stars,

(2) there are at most βn_i vertices are in the clique-stars,

(3) there are at least $(1-\beta)n_i$ vertices in one clique-star.

Next we consider the graphs in $\mathscr{B}_1 \setminus \cup_{i=1}^2 \mathcal{B}_i$. By the definition of the graphs is this set there must be an index $i \in [\mathrm{wpn}(T)]$, such that $\overline{G}[\pi'_1 \setminus S]$ does not have any of the properties (1),(2) and (3) as above, without loss of generality $i = 1$. Now we want to start using Lemma 5.2.31. Note that when we apply Lemma 5.2.31 to two clique-stars from different parts, it matters which of the clique-stars is larger. Let $\mathcal{B}_3 \subset \mathscr{B}_1 \setminus \cup_{i=1}^2 \mathcal{B}_i$ be the set of graphs where for some $i \geq 2$, $\overline{G}[\pi'_i]$ contains more than $(\log n)^2$ components from $\mathcal{S}$.

Let $G \in \mathscr{B}_1 \setminus \cup_{i=1}^3 \mathcal{B}_i$, then because $|\mathcal{S}| \geq n^\beta$ and that for each $i \geq 2$, $\overline{G}[\pi'_i]$ contains at most $(\log n)^2$ clique-stars from $\mathcal{S}$, there are at least $\frac{1}{2}n^\beta$ clique-stars from $\mathcal{S}$ in $\overline{G}[\pi'_1]$. Let $\mathcal{B}_4 \subset \mathscr{B}_1 \setminus \cup_{i=1}^3 \mathcal{B}_i$ be the set of graphs where there is an index $i \geq 2$ where both of the following properties are true for $\overline{G}[\pi'_i \setminus S]$.

(i) the largest clique-star in $\overline{G}[\pi'_i \setminus S]$ contains at most $n_i - n^{1-\beta/2}$ vertices, and

(ii) there are at least $n^{1-\beta/2}$ vertices in clique-stars in $G[\pi'_i \setminus S]$.

We first show that the size of $|\mathcal{B}_1|$ is much smaller than $|\mathscr{G}|$. Let $\mathcal{F}(\mathcal{B}_1) = (\mathcal{F}_1^1, \mathcal{F}_2^1, ..., \mathcal{F}_{\mathrm{wpn}(T)}^1)$, by the definition of $\mathcal{B}_1$ and our assumptions, for a graph $G \in \mathcal{B}_1$, $\ell_i \le \alpha n$ and $\sum_{i=1}^{\mathrm{wpn}(T)} c_i \le n^\beta$. By Lemma 5.2.26, the number of graphs in each $\mathcal{F}_i^1$, $i \in [\mathrm{wpn}(T)]$, is at most $n_i^2 \cdot 2^{2n_i} \cdot 2n^{\beta n_i} \cdot (3\alpha n)^{3\alpha n}$. Using Lemma 3.3.8 and get that the number of Π-bad graphs in this case is at most

$$\begin{aligned}
&2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot \prod_{i\in[\mathrm{wpn}(T)]} n_i^2 \cdot 2^{2n_i} \cdot 2n^{\beta n_i} \cdot (3\alpha n)^{3\alpha n} \\
&\le 2^m \cdot 2^{(b+4)n} \cdot 2^{3n} \cdot 2n^{\beta n} \cdot (3\alpha n)^{3\,\mathrm{wpn}(T)\alpha n} \\
&\le 2^m \cdot 2^{(b+7)n} \cdot 2n^{\frac{n}{10\,\mathrm{wpn}(T)^2}} \cdot n^{\frac{n}{3\,\mathrm{wpn}(T)^2}}
\end{aligned}$$

where the last inequality is due to the choice of α and β. From Observation 5.2.29 we know that the number of Π-good graphs is at least $2^m \cdot n^{\frac{n}{2\,\mathrm{wpn}(T)}}$. We compare those two bounds.

$$\begin{aligned}
\frac{|\mathcal{B}_1|}{|\mathscr{G}|} &\le \frac{2^m \cdot 2^{(b+7)n} \cdot 2n^{\frac{n}{10\,\mathrm{wpn}(T)^2}} \cdot n^{\frac{n}{3\,\mathrm{wpn}(T)^2}}}{2^m \cdot n^{\frac{n}{2\,\mathrm{wpn}(T)}}} \\
&\le \frac{2^{(b+7)n}}{n^{\frac{(\mathrm{wpn}(T)-1)n}{2\,\mathrm{wpn}(T)^2}}} = \mathrm{o}(1).
\end{aligned}$$

Next we consider graphs in $\mathscr{B}_1 \setminus \mathcal{B}_1$, by our assumptions those graphs contain at least n^β clique-stars in $\overline{G}[\pi_1'] \cup \overline{G}[\pi_2'] \cup ... \cup \overline{G}[\pi_{\mathrm{wpn}(T)}']$. Let $\mathcal{S}$ be the set of the n^β largest clique-stars in $\overline{G}[\pi_1'] \cup \overline{G}[\pi_2'] \cup ... \cup \overline{G}[\pi_{\mathrm{wpn}(T)}']$. Let S be the union of the vertices in the components in $\mathcal{S}$. Next we show that $\mathcal{B}_2 \subset \mathscr{B}_1 \setminus \mathcal{B}_1$ is much smaller than the number of Π-good graphs. Let $\mathcal{F}(\mathcal{B}_2) = (\mathcal{F}_1^2, \mathcal{F}_2^2, ..., \mathcal{F}_{\mathrm{wpn}(T)}^2)$, by the definition of $\mathcal{B}_2$, at least one of the following properties is true for $\overline{G}[\pi_i' \setminus S]$ for each $i \in [\mathrm{wpn}(T)]$,

(1) it contains at most n^β clique-stars,

(2) there are at most βn_i vertices are in clique-star,

(3) there are at least $(1-\beta)n_i$ vertices in one clique-star.

By Lemma 5.2.26, the number of ways to choose a graph on $\pi_i' \backslash S$ with property (1) is equal to $|(\mathcal{G}_2)_{n_i}^{n^\beta,*,*,\alpha n}|$ and is at most $n_i^2 \cdot 2^{2n} \cdot n^{\beta n_i} \cdot (3\alpha n)^{3\alpha n}$; The number of ways to choose a graph on $\pi_i' \backslash S$ with property (2) is equal to $|(\mathcal{G}_2)_{n_i}^{*,\beta n_i,*,\alpha n}|$ and is at most $n_i^2 \cdot 2^{3n} \cdot n^{\beta n_i} \cdot (3\alpha n)^{3\alpha n}$. Finally, the number of ways to choose a graph on $\pi_i' \backslash S$ with property (3) is equal to $|(\mathcal{G}_2)_{n_i}^{*,*,(1-\beta)n_i,\alpha n}|$ and is at most $n_i \cdot 2^{3n} \cdot (\beta n_i)^{\beta n_i} \cdot (3\alpha n)^{3\alpha n}$. In all the cases, the number of graphs on each $\overline{G}[\pi_i' \setminus S]$, $i \in [\mathrm{wpn}(T)]$, is at most

$$n_i^2 \cdot 2^{3n} \cdot n^{\beta n_i} \cdot (3\alpha n)^{3\alpha n}.$$

Therefore we can bound the number of graphs in $\mathcal{B}_2$ almost in the same way as the number of graphs in $\mathcal{B}_1$. The difference is that we need to add a factor of $n^{\beta n_i}$ because of the additional number of ways to partition the vertices in $\overline{G}[\pi_i']$ into at most n^{β} components from $\mathcal{S}$. Hence for each $i \in [\mathrm{wpn}(T)]$, $|\mathcal{F}_i^2| \leq n_i^2 \cdot 2^{3n} \cdot n^{2\beta n_i} \cdot (3\alpha n)^{3\alpha n}$, using Lemma 3.3.8 the number of graphs in $\mathcal{B}_2$ is at most

$$\begin{aligned} & 2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot \prod_{i \in [\mathrm{wpn}(T)]} 2^{3n_i} \cdot n^{2\beta n_i} \cdot (3\alpha n)^{3\alpha n} \\ & \leq 2^m \cdot 2^{(b+4)n} \cdot 2^{3n} \cdot n^{2\beta n} \cdot (3\alpha n)^{3\,\mathrm{wpn}(T)\alpha n} \\ & \leq 2^m \cdot 2^{(b+7)n} \cdot n^{\frac{n}{5\,\mathrm{wpn}(T)^2}} \cdot n^{\frac{n}{3\,\mathrm{wpn}(T)^2}}. \end{aligned}$$

In the same way as before, the number of graphs in $\mathcal{B}_2$ is much smaller than $|\mathscr{G}|$.

Let $G \in \mathscr{B}_1 \backslash \left(\cup_{i=1}^2 \mathcal{B}_1\right)$, then by the definition of this family there is an $i \in [\mathrm{wpn}(T)]$ where $\overline{G}[\pi_i' \backslash S]$ contains at least n^{β} clique-stars, at least βn_i vertices in the clique-stars and at most $(1-\beta)n_i$ vertices in one clique-star. Assume without loss of generality that $i = 1$.

We focus on the set $\mathcal{B}_3 \subset \mathscr{B}_1 \setminus \left(\cup_{i=1}^2 \mathcal{B}_1\right)$. By the definition of the family, there is an index $i \geq 2$, $\overline{G}[\pi_i']$ contains more that $(\log n)^2$ components from $\mathcal{S}$. Note that as mentioned earlier $\mathscr{B}_1 \subset \mathscr{B}$, therefore any $G \in \mathcal{B}_3$ has property (c), that is, there are no indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, where there are subgraphs J_{i_1} and J_{i_2} in $G[\pi_{i_1}']$ and $G[\pi_{i_2}']$, respectively, such that $G[V(J_{i_1}) \cup V(J_{i_2})]$ is isomorphic to P_3^*. We apply Lemma 5.2.32 to the graphs induced on $G[\pi_1']$ and $G[\pi_i' \cap S]$ and get that the number of ways to choose edges between those subgraphs is at most $2^{|\pi_1| \cdot |\pi_i \cap S|} \cdot 2^{-c\beta n \cdot (\log n)^2}$ for some constant $c > 0$. Let $\mathcal{F}(\mathcal{B}_3) = (\mathcal{F}_1^3, \mathcal{F}_2^3, ..., \mathcal{F}_{\mathrm{wpn}(T)}^3)$, we use that for each $i \in [\mathrm{wpn}(T)]$, $\mathcal{F}_i^3 \subseteq \mathcal{G}_2$, and therefore using the bound on $\mathcal{G}_2$ in Lemma 5.2.25 and Lemma 3.3.8, we can bound the number of graphs in $\mathcal{B}_3$ by,

$$\begin{aligned} & 2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot |(\mathcal{G}_2)_n| \cdot 2^{-c\beta n \cdot (\log n)^2} \\ & \leq 2^m \cdot 2^{(b+4)n} \cdot 2^{n \log n} \cdot 2^{-c\beta n \cdot (\log n)^2}. \end{aligned}$$

The values of $c > 0$ and $\beta > 0$ are constants, therefore for n large enough, the above is much smaller then 2^m and therefore much smaller than the number of Π-good graphs.

We focus on the set $\mathcal{B}_4 \subset \mathscr{B}_1 \setminus \left(\cup_{i=1}^{3}\mathcal{B}_1\right)$. Let $G \in \mathcal{B}_4$, then because for each $i \geq 2$, $\overline{G}[\pi'_i]$ contains at most $(\log n)^2$ clique-stars from $\mathcal{S}$ and $|\mathcal{S}| \geq n^\beta$, there are at least $\frac{1}{2}n^\beta$ clique-stars from $\mathcal{S}$ in $\overline{G}[\pi'_1]$. Assume that there is an index $i \geq 2$ where the largest clique-star in $\overline{G}[\pi'_i \setminus S]$ contains at most $n_i - n^{1-\beta/2}$ vertices and there are at least $n^{1-\beta/2}$ vertices in clique-stars in $\mathcal{G}[\pi'_i \setminus S]$. We apply Lemma 5.2.32 to the graphs induced on $G[\pi'_i]$ and $G[\pi'_1 \cap S]$ and get that the number of ways to choose edges between those subgraphs is at most $2^{|\pi_1 \cap S| \cdot |\pi_i|} \cdot 2^{-cn^\beta \cdot n^{1-\beta/2}}$. Let $\mathcal{F}(\mathcal{B}_4) = (\mathcal{F}_1^4, \mathcal{F}_2^4, ..., \mathcal{F}_{\text{wpn}(T)}^4)$, we use that for each $i \in [\text{wpn}(T)]$, $\mathcal{F}_i^4 \subseteq \mathcal{G}_2$, and therefore using again the bound on $\mathcal{G}_2$ in Lemma 5.2.25 and Lemma 3.3.8, we can bound the number of graphs in $\mathcal{B}_4$ by,

$$
\begin{aligned}
&2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot |(\mathcal{G}_2)_n| \cdot 2^{-cn^\beta \cdot n^{1-\beta/2}} \\
&\leq 2^m \cdot 2^{(b+4)n} \cdot 2^{n \log n} \cdot 2^{-cn^{1+\beta/2}}.
\end{aligned}
$$

The values of $c > 0$ and $\beta > 0$ are constants, therefore for n large enough, the above is much smaller then 2^m and therefore much smaller than the number of Π-good graphs.

Let $G \in \mathscr{B}_1 \setminus \left(\cup_{i=1}^{4}\mathcal{B}_1\right)$ and let $\mathcal{F}\left(\mathscr{B}_1 \setminus \left(\cup_{i=1}^{4}\mathcal{B}_1\right)\right) = (\mathcal{F}_1^5, \mathcal{F}_2^5, ..., \mathcal{F}_{\text{wpn}(T)}^5)$. From the above there is $i \in [\text{wpn}(T)]$, without loss of generality $i = 1$, such that $\overline{G}[\pi'_1]$ contains at least $\frac{1}{2}n^\beta$ components from $\mathcal{S}$, also it contains at least βn_1 vertices in the clique-stars and at most $(1-\beta)n_1$ vertices in one component. Moreover, for each $i \geq 2$, $\overline{G}[\pi'_i]$ contains at most $(\log n)^2$ components from $\mathcal{S}$, either the largest clique-star in $\overline{G}[\pi'_i \setminus S]$ contains at least $n_i - n^{1-\beta/2'}$ vertices or there are at most $n^{1-\beta/2}$ vertices in clique-stars in $\overline{G}[\pi'_i \setminus S]$. Furthermore because $c_1 \geq n^\beta$ and property (d) of graphs in $\mathcal{B}_1 \subset \mathscr{B}$, for all $i \geq 2$, $\ell_i \leq n^{1-\beta/2}$. We have that $|\mathcal{F}_1^5| \leq |(\mathcal{G}_2)_{n_1}| \leq 40^{n_1} \cdot \left(\frac{n_1}{\log\log n}\right)^{n_1}$ where the inequality is due to Lemma 5.2.25. Using Lemma 5.2.26, the number of graphs in each $\mathcal{F}_i^5$ for $i \geq 2$ is at most $n_i^2 \cdot 2^{4n_i} \cdot s_i^n \cdot 4n^{4(1-\beta/2)}$ where s_i is the number of clique-stars from $\mathcal{S}$ in $\overline{G}[\pi'_i]$. We apply again Lemma 5.2.32 to the graphs induced on $G[\pi'_1]$ and $G[\pi'_i \cap S]$ and get that the number of ways to choose edges between those subgraphs is at most $2^{|\pi_1| \cdot |\pi_i \cap S|} \cdot 2^{-c\beta n \cdot s_i}$. Using Lemma 3.3.8, we have the following bound,

$$2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot |(\mathcal{G}_2)_{n_1}| \cdot \prod_{i\geq 2} 2^{5n_i} \cdot (4n)^{4(1-\beta/2)n^{1-\beta/2}} \cdot (s_i)^n \cdot 2^{-c\beta n s_i}$$
$$\leq 2^m \cdot 2^{(b+9)n} \cdot 40^{n_1} \cdot \left(\frac{n_1}{\log\log n}\right)^{n_1} \cdot (4n)^{4\,\mathrm{wpn}(T)(1-\beta/2)n^{1-\beta/2}} \cdot 2^{n(\mathrm{wpn}(T)\log s - c\beta s)}$$

where s is the number of clique-stars in $\mathcal{S}$ which are not in $\overline{G}[\pi_1']$. We compare it to the lower bound of $2^m \cdot (\frac{n_1}{e})^{n_1}$ on the number of Π-good graphs from Observation 5.2.29. Therefore using that $n_1 \geq \frac{n}{2\,\mathrm{wpn}(T)}$,

$$\frac{|\mathscr{B}_1|}{|\mathscr{G}|} \leq \frac{2^{(b+100)n} \cdot 2^{4\,\mathrm{wpn}(T)(1-\beta/2)n^{1-\beta/2}\log n} \cdot 2^{n(\mathrm{wpn}(T)\log s - c\beta s)}}{(\log\log n)^{\frac{n}{2\,\mathrm{wpn}(T)}}}$$
$$\leq \frac{2^{(b+100)n} \cdot 2^{c'(T)n^{1-\gamma}\log n} \cdot 2^{c'(T)n}}{(\log\log n)^{\frac{n}{2\,\mathrm{wpn}(T)}}} = \mathrm{o}(1)$$

for some constants $c'(T)$ and $\gamma > 0$. Hence the number of graphs in $\mathscr{B}_1$ is much smaller than the number of Π-good graphs. □

Let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$ and let $\boldsymbol{\mathscr{B}_2 = \mathscr{B}_2(\alpha)} \subset \mathscr{B} \setminus \mathscr{B}_1$ be the subset of Π-bad graphs where there is exactly one index $i \in [\mathrm{wpn}(T)]$ such that $\ell_i \geq \alpha n$.

Lemma 5.2.41. *Let $T \in \mathcal{T}^{\mathrm{pl}}$ and let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$. The number of graphs in $\mathscr{B}_2$ is much smaller than the number of graphs in $\mathscr{G}$.*

Proof. For this proof we define the following subsets of $\mathscr{B}_2 \subset \mathscr{B} \setminus \mathscr{B}_1$. Let $G \in \mathscr{B}_2$ and let $i \in [\mathrm{wpn}(T)]$ such that $\ell_i \geq \alpha n$, assume without loss of generality that $i = 1$. Let $\beta \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^2}\right)$ and let $\mathcal{B}_1 \subset \mathscr{B}_2$ be such that $s_1 \leq 2n^\beta$.

Let $G \in \mathscr{B}_2 \setminus \mathcal{B}_1$, let $\mathcal{S}_1$ be the set of n^β largest clique-stars in $\overline{G}[\pi_1']$ and let S_1 be the union of vertices in $\mathcal{S}_1$. Let $\mathcal{B}_2 \subset \mathscr{B}_2 \setminus \mathcal{B}_1$ be the set of graphs where one the following is true for $\overline{G}[\pi_1' \setminus S_1]$ (note that we consider the subgraph on the vertex set without the vertices in the largest clique-stars),

(a) it contains at most n^β clique-stars,

(b) there are at most βn_1 vertices are in clique-stars,

(c) there are at least $(1-\beta)n_1$ vertices in one clique-star.

We describe how we use those families. Let $G \in \mathscr{B}_2 \subset \mathscr{B}$, by our assumptions $\ell_1 \geq \alpha n$ and therefore we can find at least αn disjoint P_3 in $G[\pi'_1]$. Then by part (d) of the definition of $\mathscr{B}$ this implies that for each $i \geq 2$, $c_i \leq (\log n)^2$.

We consider now the set $\mathcal{B}_1 \subset \mathscr{B}_2$, let $\mathcal{F}(\mathcal{B}_1) = (\mathcal{F}^1_1, \mathcal{F}^1_2, ..., \mathcal{F}^1_{\operatorname{wpn}(T)})$. By the definition of $\mathcal{B}_1$ in each $G \in \mathcal{B}_1$, $c_1 \leq n^\beta$, therefore by Lemma 5.2.27, $|\mathcal{F}^1_1| \leq 2^{n_1} \cdot n^{\beta n_1} \cdot n_1^{\frac{n_1}{2}}$. In each $G \in \mathcal{B}_1$, for each $i \geq 2$, $c_i \leq (\log n)^2$ and $\ell_i \leq \alpha n$, therefore using Lemma 5.2.26, for each $i \geq 2$, $|\mathcal{F}^1_i| \leq n_i^2 \cdot 2^{2n_i} \cdot (\log n)^{2n_i} \cdot (3\alpha n)^{3\alpha n}$. Using Lemma 3.3.8 the number of graphs in $\mathcal{B}_1$ is at most

$$\begin{aligned}
&2^m \cdot 2^{(b+3)n} \cdot 2^{-c(T)ny} \cdot 2^{n_1} \cdot (2n)^{\beta n_1} \cdot n_1^{\frac{n_1}{2}} \cdot \prod_{i \geq 2} n_i^2 \cdot 2^{2n_i} \cdot (\log n)^{2n_i} \cdot (3\alpha n)^{3\alpha n} \\
&\leq 2^m \cdot 2^{(b+8)n} \cdot n^{\beta n_1} \cdot n_1^{\frac{n_1}{2}} \cdot (\log n)^{2n} \cdot (3\alpha n)^{3\operatorname{wpn}(T)\alpha n}
\end{aligned}$$

we compare it to the lower bound in Observation 5.2.29.

$$\begin{aligned}
\frac{|\mathcal{B}_1|}{|\mathscr{G}|} &\leq \frac{2^m \cdot 2^{(b+8)n} \cdot n^{\beta n_1} \cdot n_1^{\frac{n_1}{2}} \cdot (\log n)^{2n} \cdot (3\alpha n)^{3\operatorname{wpn}(T)\alpha n}}{2^m \cdot \left(\frac{n_1}{e}\right)^{\frac{n_1}{2}} \cdot \left(\frac{n_2}{e}\right)^{\frac{n_2}{2}}} \\
&\leq \frac{2^{(b+9)n} \cdot n^{\beta n_1} \cdot (\log n)^{2n} \cdot n^{3\operatorname{wpn}(T)\alpha n}}{(n_2)^{\frac{n_2}{2}}} \\
&\leq \frac{2^{(b+10)n} \cdot n^{\frac{n}{\operatorname{wpn}(T)^2}} \cdot (\log n)^{2n} \cdot n^{\frac{n}{\operatorname{wpn}(T)^2}}}{n^{\frac{n}{4\operatorname{wpn}(T)}}} = \mathrm{o}(1).
\end{aligned}$$

Next we consider graphs in $\mathcal{B}_2$, let $\mathcal{F}(\mathcal{B}_2) = (\mathcal{F}^2_1, \mathcal{F}^2_2, ..., \mathcal{F}^2_{\operatorname{wpn}(T)})$. Similarly to the arguments in Lemma 5.2.40, using Lemma 5.2.27, $|\mathcal{F}^2_1| \leq 2^{4n} \cdot n^{2\beta n_1} \cdot n_1^{\frac{n_1}{2}}$. Using almost the same computation as above, we get that the number of graphs in $\mathcal{B}_2$ is much smaller than the number of graphs in $\mathscr{G}$.

Let $G \in \mathscr{B}_2 \setminus \cup_{i=1}^2 \mathcal{B}_i$, then $\ell_1 \geq \alpha n$, $c_1 \geq n^\beta$, $\overline{G}[\pi'_1 \setminus S_1]$ contains at least n^β clique-stars, at least βn_1 vertices in those components and no component contains at least $(1-\beta)n_1$ vertices. Moreover, by property (d) of the graphs in $\mathscr{B}$, for each $i \geq 2$, $c_i \leq (\log n)^2$ and $\ell_i \leq n^{1-\beta/2}$.

Let t be the size of the largest component in $\overline{G}[\pi'_1 \setminus S_1]$ and let $\mathcal{K}_1$ be the clique-stars of size at most t in $\overline{G}[\pi'_1 \setminus S_1]$ (and so in $\overline{G}[\pi'_1]$). For $i \geq 2$, let $\mathcal{S}_i$ be the clique-stars in $\overline{G}[\pi'_i]$ of size at least t and let $\mathcal{K}_i$ be the clique-stars in $\overline{G}[\pi'_i]$ of size at most t. For $i \geq 2$, let $|\mathcal{S}_i| = s_i$ and let S_i be union of vertices in all clique-stars in $\mathcal{S}_i$ in $\overline{G}[\pi'_i]$. Similarly, let K_i be the union of vertices in all clique-stars in $\mathcal{K}_i$ in $\overline{G}[\pi'_i]$ and $k_i = |K_i|$.

Let $\mathcal{F}(\mathscr{B}_2 \setminus \cup_{i=1}^2 \mathcal{B}_i) = (\mathcal{F}_1^3, \mathcal{F}_2^3, ..., \mathcal{F}_{\text{wpn}(T)}^3)$. Let $i \geq 2$ and let $n''_i = |\pi_i \setminus S_i|$. The number of graphs in $\mathcal{F}_i^3$ such that one of the components in $\mathcal{K}_i$ contains at least $n''_i - n^\beta$ vertices is at most $s_i^n \cdot (\mathcal{G}_2)_{n''_i}^{*,*,n''_i - n^\beta, n^{1-\beta/2}} \leq s_i^n \cdot 2^{4n_i} \cdot n^{\beta n^\beta} \cdot (3n)^{3(1-\beta/2)n^{(1-\beta/2)}}$ where the bound is by Lemma 5.2.26. Let $U \subset [\text{wpn}(T)] \setminus \{1\}$ be the set of indices such that for each $i \in U$, one of the components in $\mathcal{K}_i$ contains at least $n''_i - n^\beta$.

Let $i \notin U \cup \{1\}$, then we can bound the number of graphs in $\mathcal{F}_i^3$ by $2^{4n_i} \cdot k_i^{k_i} \cdot s_i^n \cdot (3n)^{3(1-\beta/2)n^{1-\beta/2}}$. There are at most 2^{4n_i} ways to choose the elements in S_i, K_i, components of size 2 in $\overline{G}[\pi_i]$, and the centres of the clique-stars. There are at most $k_i^{k_i}$ ways to partition the vertices in K_i into the different clique-stars. There are at most s_i^n ways to partition the vertices into the s_i different components on S_i. Finally, there are at most $(3n)^{3(1-\beta/2)n^{1-\beta/2}}$ ways to partition the $n^{1-\beta/2}$ non-edges in $\overline{G}[\pi_i]$.

Finally, we apply Lemma 5.2.32 to $G[\pi'_1 \setminus S_1]$ and $G[S_i]$ for each $i \geq 2$. Moreover, we apply Lemma 5.2.32 to $G[S_1]$ and each $G[K_i]$, such that $i \notin U \cup \{1\}$. Then using also Lemma 3.3.8, we can bound the number of graphs in $\mathscr{B}_2 \setminus \cup_{i=1}^2 \mathcal{B}_i$ by

$$
\begin{aligned}
&2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot |(\mathcal{G}_2)_{n_1}| \cdot \prod_{i \in U} s_i^n \cdot 2^{4n_i} \cdot n^{\beta n^\beta} \cdot (3n)^{3(1-\beta/2)n^{(1-\beta/2)}} \cdot 2^{-c\beta n s_i} \cdot \\
&\qquad\qquad\qquad \cdot \prod_{i \notin U \cup \{1\}} 2^{4n_i} \cdot k_i^{k_i} \cdot s_i^n \cdot (3n)^{3(1-\beta/2)n^{1-\beta/2}} \cdot 2^{-ck_i n^\beta} \cdot 2^{-c\beta n s_i} \\
&\leq 2^m \cdot 2^{(b+8)n} \cdot |(\mathcal{G}_2)_{n_1}| \cdot s^{\text{wpn}(T)n} \cdot n^{\beta n^\beta} \cdot n^k \cdot n^{5\,\text{wpn}(T)(1-\beta/2)n^{1-\beta/2}} \cdot 2^{-ckn^\beta} \cdot 2^{-c\beta n s} \\
&\leq 2^m \cdot 2^{(b+8)n} \cdot 40^{n_1} \cdot \left(\frac{n_1}{\log \log n_1} \right)^{n_1} \cdot 2^{10\,\text{wpn}(T)(1-\beta/2)n^{1-\beta/2} \log n} \cdot 2^{k(\log n - cn^\beta)} \cdot 2^{n(\text{wpn}(T) \log s - c\beta s)}
\end{aligned}
$$

where $s = \sum_{i \geq 2} s_i$ and $k = \sum_{i \notin U \cup \{1\}} k_i$ and $c > 0$. We compare it to the lower bound of $2^m \cdot (\frac{n_1}{e})^{n_1/2} \cdot (\frac{n_2}{e})^{n_2/2}$ on the number of Π-good graphs $\mathscr{G}$ from Observation 5.2.29. Therefore using that Π is a $\rho/4$-almost equal partition and every two parts of the partition can differ on at $2n^{1-\rho/4}$ elements,

$$\begin{aligned}\frac{|\mathscr{B}_2|}{|\mathscr{G}|} &\leq \frac{2^{(b+15)n} \cdot 2^{10\,\mathrm{wpn}(T)(1-\beta/2)n^{1-\beta/2}\log n} \cdot 2^{k(\log n - cn^{\beta})} \cdot 2^{n(\mathrm{wpn}(T)\log s - c\beta s)}}{(\log\log n)^{\frac{n}{4\,\mathrm{wpn}(T)}}} \\ &\leq \frac{2^{(b+15)n} \cdot 2^{c'(T)n^{1-\gamma}\log n} \cdot 2^{c'(T)n}}{(\log\log n)^{\frac{n}{4\,\mathrm{wpn}(T)}}} = \mathrm{o}(1),\end{aligned}$$

where $\gamma > 0$ and $c(T)$ a constant which depends only on T. Therefore the number of graphs in $\mathscr{B}_2$ is much smaller than the number of graphs in $\mathscr{G}$ ☐

Let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$, and $G \in \mathscr{B} \setminus (\mathscr{B}_1 \cup \mathscr{B}_2)$, due to property (d) of graphs in $\mathscr{B}$, we can deduce the following.

Observation 5.2.42. *Let $T \in \mathcal{T}^{\mathrm{pl}}$ and let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$ and $G \in \mathscr{B} \setminus (\mathscr{B}_1 \cup \mathscr{B}_2)$. Let $i_1 \neq i_2$ be the indices such that $\ell_{i_1} \geq \alpha n$ and $\ell_{i_2} \geq \alpha n$, then $c_i \leq (\log n)^2$ for all $i \in [\mathrm{wpn}(T)]$.*

Trees in $\mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$: Let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$ and let $G \in \mathscr{B} \setminus (\mathscr{B}_1 \cup \mathscr{B}_2)$. Let $W = W(G) \subset [n]$ be a minimal set such that for each $i \in [\mathrm{wpn}(T)]$ $G[\pi_i \setminus W(G)] \in \mathcal{G}_1$, that is each $G[\pi_i \setminus W(G)]$ is a disjoint union of edges and singleton vertices.

Lemma 5.2.43. *Let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$ and let $G \in \mathscr{B} \setminus (\mathscr{B}_1 \cup \mathscr{B}_2)$, then $|W(G)| \leq 2\,\mathrm{wpn}(T) \cdot (\log n)^2$.*

Proof. Let C_i be the centres of the clique-stars in $\overline{G}[\pi_i']$, $i \in [\mathrm{wpn}(T)]$. Note that $G[\pi_i' \setminus C_i]$ is in $\mathcal{G}_1$. Therefore we can define $W(G)$ to be $\left(\cup_{i=1}^{\mathrm{wpn}(T)} C_i\right) \cup Y$. By property (b) in the definition of the set $\mathscr{B}$, $y \leq C(T)\log n$. By Observation 5.2.42, there are at most $(\log n)^2$ clique-stars, and therefore $|C_i| \leq (\log n)^2$. ☐

Let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$, $\gamma \in \left(0, \frac{1}{2\,\mathrm{wpn}(T)}\right)$ and let $\boldsymbol{\mathscr{B}_3 = \mathscr{B}_3(\gamma)} \subset \mathscr{B} \setminus (\mathscr{B}_1 \cup \mathscr{B}_2)$ be the set of graphs where there is $i \in [\mathrm{wpn}(T)]$ such that $\overline{G}[\pi_i \setminus W(G)]$ contains at least γn singleton components.

Lemma 5.2.44. *Let $T \in \mathcal{T}^{\mathrm{pl}}_{\mathrm{star}}$ and let $\alpha \in \left(0, \frac{1}{10\,\mathrm{wpn}(T)^3}\right)$, $\gamma \in \left(0, \frac{1}{2\,\mathrm{wpn}(T)}\right)$. The number of graphs in $\mathscr{B}_3$ is much smaller than the number of graphs in $\mathscr{G}$.*

Proof. Let $\mathcal{F}(\mathscr{B}_3) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$. Let $i \in [\text{wpn}(T)]$ be such $\overline{G}[\pi_i \setminus W]$ contains at least γn singleton components, then by Lemma 5.2.24, $|\mathcal{F}_i| \leq 2^{n_i} \cdot (n_i - \gamma n)^{\frac{n_i - \gamma n}{2}} \cdot (\log n)^{2n_i}$. The number of graphs in each $\mathcal{F}_j$ for $j \in [\text{wpn}(T)] \setminus \{i\}$ is at most $n_j^{\frac{n_j}{2}} \cdot (\log n)^{2n_j}$. We compare it to the lower bound on the number of Π-good graphs from Observation 5.2.29 and get the following.

$$\frac{|\mathscr{B}_3|}{|\mathscr{G}|} \leq \frac{2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot 2^{n_i} \cdot (n_i - \gamma n)^{\frac{n_i - \gamma n}{2}} \cdot (\log n)^{2n_i} \cdot \prod_{j \neq i} n_j^{\frac{n_j}{2}} \cdot (\log n)^{2n_j}}{2^m \cdot \left(\frac{n}{e}\right)^{\frac{n}{2}}}$$

$$\leq \frac{2^{(b+8)n} \cdot (\log n)^{2n} \cdot n^{\frac{n}{2}} \cdot n_i^{\frac{-\gamma n}{2}}}{n^{\frac{n}{2}}}$$

$$\leq 2^{(b+8)n} \cdot (\log n)^{2n} \cdot \left(\frac{n}{2\,\text{wpn}(T)}\right)^{-\frac{\gamma n}{2}} = \text{o}(1),$$

where the last inequality due to the fact that $n_i \geq \frac{n}{2\,\text{wpn}(T)}$. Hence we get the required result. ☐

Let $\alpha \in \left(0, \frac{1}{10\,\text{wpn}(T)^3}\right)$, $\gamma \in \left(0, \frac{1}{2^{4t^2}\,\text{wpn}(T)^3}\right)$, we focus on graphs in $\mathscr{B} \setminus \cup_{i=1}^{3} \mathscr{B}_i$. Let $\boldsymbol{\mathcal{F}'(\mathscr{B} \setminus \cup_{i=1}^{3} \mathscr{B}_i) = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\text{wpn}(T)})}$, that is each $\mathcal{F}'_i \subseteq \mathcal{G}_1$ and for each $G \in \mathcal{F}'_i$ and $i \in [\text{wpn}(T)]$, $\overline{G}$ contains at most γn singleton components.

Lemma 5.2.45. *Let $\varepsilon, \varepsilon' > 0$ to be the constants defined earlier in the subsection. The sequence $\mathcal{F}'$ is (τ, λ, k)-ordinary $P(T)$-free subsequence for any $\lambda \in [0, \frac{\min\{\varepsilon, \varepsilon'\}}{8})$, $\tau = \frac{1}{w^2}$, and any $k \in \mathbb{N}$.*

Proof. We check properties (a) and (b) of the definition of (τ, λ, k)-ordinary $P(T)$-free subsequence. Let $Q = (G_1, G_2, ..., G_{\text{wpn}(T)})$ be a sequence of graphs such that $G_i \in \mathcal{F}'_i$ for each $i \in [\text{wpn}(T)]$.

First we check property (a). Let $i \in [\text{wpn}(T)]$, the graph $G_i \in \mathcal{G}_1$, therefore for each $v \in V(G_i)$, $\overline{\deg}_{G_i}(v) \in \{0, 1\}$. Hence v is (τ, G_i)-trivial for any $\tau > 0$. Let $\tau = \frac{1}{w^2}$.

Now we check property (b). Let $w' \notin \cup_{i \in [\text{wpn}(T)]} V(G_i)$ be such for each $i \in [\text{wpn}(T)]$, $G_i \cup \{w'\} \notin \mathcal{F}'_i$. Let $\ell_i(w') = \min\{|N(w') \cap V(G_i)|, |\overline{N}(w') \cap V(G_i)|\}$, $i \in [\text{wpn}(T)]$, let $\ell(w') = \min_{i \in [\text{wpn}(T)]} \ell_i(w')$ and let $i \in [\text{wpn}(T)]$ be such that $\ell(w') = \ell_i(w')$, then for each $j \in [\text{wpn}(T))] \setminus \{i\}$, $\ell_j(w') \geq \tau n$, that is w' is not (τ, G_j)-trivial.

Let $\mathcal{K}$ be a maximal collection of disjoint edges or non-edges in $V(G_i)$ such that for each $K \in \mathcal{K}$, $K \cup \{w'\}$ is isomorphic to either S_3 or $\overline{P_3}$. The set $\mathcal{K} \neq \emptyset$ because $G_i \cup \{w'\} \notin \mathcal{F}'_i$. Assume that $|\mathcal{K}| \geq \gamma n$, then there are at least $|\mathcal{K}| - \gamma n$ disjoint non-edges in $\overline{N}(w') \cap V(G_i)$, let $\mathcal{N} \subseteq \mathcal{K}$ be this set of non-edges. Let $N \in \mathcal{N}$, then $N \cup \{w'\}$ is isomorphic to either $\overline{P_3}$ or S_3. By Lemma 5.2.19, it is possible to find in G_1 a set $\mathcal{P}$ of disjoint copies of P_3 such that $|\mathcal{P}| \geq \frac{\tau n - \gamma n}{4}$ and either (i) w' is adjacent to exactly one end (and not other vertices) of each copy of P_3 in $\mathcal{P}$ or (ii) w' is adjacent to the centre vertex (an not other vertices) of each copy of P_3 in $\mathcal{P}$. Assume without loss of generality that $\mathcal{N}$ contains a set $\mathcal{N}'$ of at least $\frac{|\mathcal{N}|}{2}$ non-edges such that each such non-edge together with w' induces a copy of $\overline{P_3}$. Note that in each such $\overline{P_3}$, w' is an end of an edge. By our assumptions, for each $j \in [\text{wpn}(T)] \setminus \{1, i\}$, w' is not (τ, G_j)-trivial, therefore for each of the possible edge arrangement between a vertex and an edge, G_j contains a set $\mathcal{H}_j$ of at least $\frac{\tau n}{4}$ disjoint edges, so w' has this edge arrangement.

Let $P(T) = (P_3, H_2, ..., H_{\text{wpn}(T)})$ be a partition of T where $H_i = \overline{P_3}$ and each H_j for $j \in [\text{wpn}(T)] \setminus \{1, i\}$ is an edge, let w be the centre of the $H_i = \overline{P_3}$. The tree T has such a partition by Lemma 5.2.2 and Observation 5.2.7. By part (e) of Observation 5.2.30 and the above, w' is adjacent to each copy P_3 in $\mathcal{P} = \mathcal{H}_1$ as w to P_3 in the partition $P(T)$. The collection of sets $(\mathcal{P}, \mathcal{H}_2, .., \mathcal{H}_{i-1}, \mathcal{N}', \mathcal{H}_{i+1}, ..., \mathcal{H}_{\text{wpn}(T)})$ is a (P, w', i)-form in Q with property (1) of (b). □

Now we complete the proof of the exact structure for all trees $T \in \mathcal{T}^{\text{pl}}_{\text{star}}$.

Proof of Theorem 1.1.9 for trees $T \in \mathcal{T}^{\text{pl}}_{\text{star}}$. From Observation 5.2.11, to show that the number of bad graphs is much smaller than the number of good graphs it is enough to show that the number of Π-bad graphs is much smaller than the number of Π-good graphs. From Lemmas 5.2.38, 5.2.39, 5.2.40, 5.2.41, Observation 5.2.42, and Lemmas 5.2.43, 5.2.44, we have the following. Almost all Π-bad graphs G have a set $W(G) \subset V(G)$ such that $|W(G)| \leq 2\,\text{wpn}(T) \cdot (\log n)^2$, for each $i \in [\text{wpn}(T)]$, $G[\pi_i \setminus W(G)] \in \mathcal{G}_1$, and for each $i \in [\text{wpn}(T)]$, $\overline{G}[\pi_i \setminus W(G)]$ contains at most γn singleton vertices for any fixed $\gamma \in \left(0, \frac{1}{2\,\text{wpn}(T)}\right)$, let $\mathscr{D}$ be this collection of graphs. Let $\gamma \in \left(0, \frac{1}{2^{4t^2}\,\text{wpn}(T)^3}\right)$ and let $\mathcal{F}' = \mathcal{F}(\mathscr{D})$, then by Lemma 5.2.45,

$\mathcal{F}'$ is (τ, λ, k)-ordinary $P(T)$-free subsequence for $\tau = \frac{1}{w^2}$, any λ small enough and any $k \in \mathbb{N}$. Finally, by Lemma 3.3.14, the number of the remaining Π-bad graphs is much smaller than the number of Π-good graphs. $\square$

Trees in $\mathcal{T}^{\text{pl}} \backslash \mathcal{T}^{\text{pl}}_{\text{star}}$: Let $\mathscr{B}_{\mathbf{3}} \subset \mathscr{B} \backslash (\mathscr{B}_1 \cup \mathscr{B}_2)$ be the set of graphs where $\ell_{i_1} \geq \alpha n$ and $\ell_{i_2} \geq \alpha n$ for some $i_1 \neq i_2 \in [\text{wpn}(T)]$ and there is $i \in [\text{wpn}(T)] \setminus \{i_1, i_2\}$, such that $G[\pi'_i]$ contains at least $(\log n)^2$ disjoint non-edges. This is a direct corollary to Corollary 5.2.37

Corollary 5.2.46. *Let $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$ and let $\alpha \in \left(0, \frac{1}{10 \, \text{wpn}(T)^3}\right)$. The number of graphs in $\mathscr{B}_3$ is much smaller than the number of graphs in $\mathscr{G}$.*

Let $G \in \mathscr{B} \setminus \left(\cup_{i=1}^{3} \mathscr{B}_i\right)$ let $i_1 \neq i_2 \in [\text{wpn}(T)]$ so $\ell_{i_1} \geq \alpha n$ and $\ell_{i_2} \geq \alpha n$, assume without loss of generality that $i_1 = 1$ and $i_2 = 2$. Let $W = W(G) \subset [n]$ be a minimal set such that $G[\pi_i \setminus W]$ and $G[\pi_2 \setminus W]$ are in $\mathcal{G}_1$ and $G[\pi_i \setminus W]$ for each $i \in [\text{wpn}(T)] \setminus \{1, 2\}$ is a clique.

Lemma 5.2.47. *Let $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$ and let $\alpha \in \left(0, \frac{1}{10 \, \text{wpn}(T)^3}\right)$. Let $G \in \mathscr{B} \setminus \left(\cup_{i=1}^{3} \mathscr{B}_i\right)$, then $|W(G)| \leq 2 \, \text{wpn}(T) \cdot (\log n)^2$.*

Proof. Let C_1 and C_2 be the centres of the clique-stars in $\overline{G}[\pi'_1]$ and $\overline{G}[\pi'_2]$, and let C_i for $i \in [\text{wpn}(T)] \setminus \{1, 2\}$ to be the union of the centres of clique-stars and the non-edges in $\overline{G}[\pi'_i]$. Note that $G[\pi'_1 \setminus C_1]$ and $G[\pi'_2 \setminus C_2]$ are in $\mathcal{G}_1$, and $G[\pi_i \setminus W]$ is a clique for each $i \in [\text{wpn}(T)] \setminus \{1, 2\}$. Therefore we can define W to be $\left(\cup_{i=1}^{\text{wpn}(T)} C_i\right) \cup Y$. By property (b) in the definition of the set $\mathscr{B}$, $y \leq C(T) \log n$. By Observation 5.2.42 and Corollary 5.2.46, for all $i \in [\text{wpn}(T)]$ $|C_i| \leq (\log n)^2$. Hence we get the required bound. $\square$

Let $\alpha \in \left(0, \frac{1}{10 \, \text{wpn}(T)^3}\right)$ and let $\gamma \in \left(0, \frac{1}{2 \, \text{wpn}(T)}\right)$. Let $\mathscr{B}_4 \subset \mathscr{B} \setminus \left(\cup_{i=1}^{3} \mathscr{B}_i\right)$ to be the set of graphs where there either $\overline{G}[\pi_1 \setminus W]$ or $\overline{G}[\pi_2 \setminus W]$ contains at least γn singleton components.

Lemma 5.2.48. *Let $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$ and let $\alpha \in \left(0, \frac{1}{10 \, \text{wpn}(T)^3}\right)$ and let $\gamma \in \left(0, \frac{1}{2 \, \text{wpn}(T)}\right)$. The number of graphs in $\mathscr{B}_4$ is much smaller than the number of graphs in $\mathscr{G}$.*

Proof. Let $j \in \{1, 2\}$ be such $\overline{G}[\pi_j \setminus W]$ contains at least γn singleton components, then by Lemma 5.2.24, the number of graphs in $G[\pi'_j \backslash W]$ is at most $2^{n_j} \cdot \left(n_j - \gamma n\right)^{\frac{n_j - \gamma n}{2}}$. We assume without loss of generality that $j = 1$. Therefore similarly to the proof of Lemma 5.2.44,

$$\begin{aligned}\frac{|\mathscr{B}_4|}{|\mathscr{G}|} &\leq \frac{2^m \cdot 2^{(b+4)n} \cdot 2^{-c(T)ny} \cdot 2^{n_1} \cdot (n_1 - \gamma n)^{\frac{n_1 - \gamma n}{2}} \cdot n_2^{\frac{n_2}{2}} \cdot (\log n)^{2n}}{2^m \cdot \left(\frac{n_1}{e}\right)^{\frac{n_1}{2}} \cdot \left(\frac{n_2}{e}\right)^{\frac{n_2}{2}}} \\ &\leq \frac{2^{(b+8)n} \cdot (\log n)^{2n} \cdot n_1^{\frac{n_1}{2}} \cdot n_i^{\frac{-\gamma n}{2}}}{n_1^{\frac{n_1}{2}}} \\ &\leq 2^{(b+8)n} \cdot (\log n)^{2n} \cdot \left(\frac{n}{2\operatorname{wpn}(T)}\right)^{-\frac{\gamma n}{2}} = \mathrm{o}(1),\end{aligned}$$

where the last inequality due to the fact that $n_i \geq \frac{n}{2\operatorname{wpn}(T)}$. Hence we get the required result. □

Let $\alpha \in \left(0, \frac{1}{10\operatorname{wpn}(T)^3}\right)$, $\gamma \in \left(0, \frac{1}{2^{4t^2}\operatorname{wpn}(T)^3}\right)$, we focus on graphs in $\mathscr{B} \setminus \cup_{i=1}^{4} \mathscr{B}_i$. Let $\boldsymbol{\mathcal{F}'(\mathscr{B} \setminus \cup_{i=1}^{4} \mathscr{B}_i) = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\operatorname{wpn}(T)})}$, that $\mathcal{F}'_1 \subseteq \mathcal{G}_1$, $\mathcal{F}'_2 \subseteq \mathcal{G}_1$ and for each $G \in \mathcal{F}'_1 \cup \mathcal{F}'_2$, $\overline{G}$ contains at most γn singleton components. For each $i \in [\operatorname{wpn}(T)] \setminus \{1, 2\}$, $\mathcal{F}'_i$ contains a clique.

Lemma 5.2.49. *Let $\varepsilon, \varepsilon' > 0$ to be the constants defined earlier in the subsection. The sequence $\mathcal{F}'$ is (τ, λ, k)-ordinary $P(T)$-free subsequence for any $\lambda \in [0, \frac{\min\{\varepsilon, \varepsilon'\}}{8})$, $\tau = \frac{1}{w^2}$, and any $k \in \mathbb{N}$.*

Proof. We check properties (a) and (b) of the definition of (τ, λ, k)-ordinary $P(T)$-free subsequence. Let $Q = (G_1, G_2, ..., G_{\operatorname{wpn}(T)})$ be a sequence of graphs such that $G_i \in \mathcal{F}'_i$ for each $i \in [\operatorname{wpn}(T)]$.

First we check property (a). Let $i \in \{1, 2\}$, the graph $G_i \in \mathcal{G}_1$, therefore for each $v \in V(G_i)$, $\overline{\deg}_{G_i}(v) \in \{0, 1\}$. Let $i \geq 3$, then G_i is a clique, therefore for each $v \in V(G_i)$, $\overline{\deg}_{G_i}(v) = 0$. Hence v is (τ, G_i)-trivial for any $\tau > 0$. Let $\tau = \frac{1}{w^2}$.

Now we check property (b). Let $w' \notin \cup_{i \in [\operatorname{wpn}(T)]} V(G_i)$ be such for each $i \in [\operatorname{wpn}(T)]$, $G_i \cup \{w'\} \notin \mathcal{F}'_i$. Let $\ell_i(w') = \min\{|N(w') \cap V(G_i)|, |\overline{N}(w') \cap V(G_i)|\}$, $i \in [\operatorname{wpn}(T)]$, let $\ell(w') = \min_{i \in [\operatorname{wpn}(T)]} \ell_i(w')$ and let $i \in [\operatorname{wpn}(T)]$ be such that $\ell(w') = \ell_i(w')$, then for each $j \in [\operatorname{wpn}(T)] \setminus \{i\}$, $\ell_j(w') \geq \tau n$, that is w' is not (τ, G_j)-trivial.

If $i \in \{1, 2\}$, then the proof is the same as in Lemma 5.2.45. Therefore we consider the case where $i \geq 3$. It is possible to find in G_i a set $\mathcal{N}$ of disjoint edges such that $|\mathcal{N}| \geq \frac{\ell(w')}{2}$ and each edge $N \in \mathcal{N}$ together with w' induce either S_2, $\overline{P_3}$ or P_3.

From Lemma 5.2.19 in both G_1 and G_2, it is possible to find a set $\mathcal{P}$ of disjoint copies of P_3 such that $|\mathcal{P}| \geq \frac{\tau n - \gamma n}{4}$ and either (i) w' is adjacent to exactly one end (and not other vertices) of each copy of P_3 in $\mathcal{P}$ or (ii) w' is not adjacent to any of the vertices of each copy of P_3 in $\mathcal{P}$. Note that conditions (i) and (ii) also imply that there is a set $\mathcal{P}'$ of disjoint non-edges such that w' is either adjacent to exactly one end or not adjacent to any of those non-edges and $|\mathcal{P}'| = |\mathcal{P}|$.

We consider 3 different cases: (a) G_1 and G_2 contain sets $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively, such that w' is adjacent to each P_3 in $\mathcal{P}_1$ and $\mathcal{P}_2$ as in case (i), (b) G_1 and G_2 contain sets $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively, such that w' is adjacent to each P_3 in $\mathcal{P}_1$ and $\mathcal{P}_2$ as in case (ii) and (c) G_1 and G_2 contain sets $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively, such that without loss of generality w' is adjacent to each P_3 in $\mathcal{P}_1$ as in case (i) and to each P_3 in $\mathcal{P}_2$ as in case (ii).

Note that by Lemma 5.2.4, each tree $T \in \mathcal{T}^{\text{pl}} \setminus \mathcal{T}^{\text{pl}}_{\text{star}}$ has a partition into P_4^* and $\text{wpn}(T) - 3$ edges. In case (a), we set $\mathcal{N}$ to be the maximal collection of disjoint edges in $\overline{N}(w') \cap V(G_i)$ and $\mathcal{P}'_1$ to be a collection of non-edges that we get from $\mathcal{P}_1$. By our assumptions, for each $j \in [\text{wpn}(T)] \setminus \{i\}$, w' is not (τ, G_j)-trivial, therefore for each of the possible edge arrangement between a vertex and an edge, G_j contains a set $\mathcal{H}_j$ of at least $\frac{\tau n}{4}$ disjoint edges, so w' has this edge arrangement with each edge in $\mathcal{H}_j$. Let $P_1 = (S_2, P_3, ..., H_{\text{wpn}(T)})$ where $H_i = \overline{P_3}$ and for each $j \in [\text{wpn}(T)] \setminus \{1, 2, i\}$, H_j is an edge. As shown in Observation 5.2.20, part a, there is a partition of P_4^* such that one of the vertices $w \in V(P_4^*)$ is adjacent to one end of a non-edge, one end of a P_3 and induce a $\overline{P_3}$ with the remaining edge. Therefore the collection of sets $(\mathcal{P}'_1, \mathcal{P}_2, .., \mathcal{H}_{i-1}, \mathcal{N}, \mathcal{H}_{i+1}, ..., \mathcal{H}_{\text{wpn}(T)})$ is a (P_1, w', i)-form with property (2) of (b).

In case (b), we set $\mathcal{N}$ to be the maximal collection of disjoint edges in G_i such that w' is adjacent to exactly one end of each such edge. Note that for any $N \in \mathcal{N}$, $N \cup \{w'\}$ induces a P_3. Let $\mathcal{P}'_1$ to be a collection of non-edges that we get from $\mathcal{P}_1$. By our assumptions, for each $j \in [\text{wpn}(T)] \setminus \{i\}$, w' is not (τ, G_j)-trivial, therefore as before for each of the possible edge arrangements between a vertex and an edge, G_j contains a set $\mathcal{H}_j$ of at least $\frac{\tau n}{4}$ disjoint edges, so w' has this edge arrangement with every edge in $\mathcal{H}_j$. Let $P_2 = (S_2, P_3, ..., H_{\mathrm{wpn}(T)})$ where $H_i = P_3$ and for each $j \in [\mathrm{wpn}(T)] \setminus \{1, 2, i\}$, H_j is an edge. As shown in Observation 5.2.20, part b, there is a partition of P_4^* such that one of the vertices $w \in V(P_4^*)$ is not adjacent to the non-edge, and not adjacent to P_3 and induce a P_3 with the remaining edge. Therefore the collection of sets $(\mathcal{P}_1', \mathcal{P}_2, .., \mathcal{H}_{i-1}, \mathcal{N}, \mathcal{H}_{i+1}, ..., \mathcal{H}_{\mathrm{wpn}(T)})$ is a (P_2, w', i)-form with property (2) of (b).

In case (c), we set $\mathcal{N} = \overline{N}(w') \cap V(G_1)$. Note that for any $N \in \mathcal{N}$, $N \cup \{w'\}$ is a non-edge. By our assumptions, for each $j \in [\mathrm{wpn}(T)] \setminus \{i\}$, w' is not (τ, G_j)-trivial, therefore as before for each of the possible edge arrangements between a vertex and an edge, G_j contains a set $\mathcal{H}_j$ of at least $\frac{\tau n}{4}$ disjoint edges, so w' has this edge arrangement with every edge in $\mathcal{H}_j$. Let $P_3 = (P_2, P_3, ..., H_{\mathrm{wpn}(T)})$ where $H_i = S_2$ and for each $j \in [\mathrm{wpn}(T)] \setminus \{1, 2, i\}$, H_j is an edge. As shown in Observation 5.2.20, part c, there is a partition of P_4^* such that one of the vertices $w \in V(P_4^*)$ is not adjacent to one of the P_3, and adjacent to one of the ends of the other P_3 and induce a S_2 with the remaining vertex. Therefore the collection of sets $(\mathcal{P}_1', \mathcal{P}_2, .., \mathcal{H}_{i-1}, \mathcal{N}, \mathcal{H}_{i+1}, ..., \mathcal{H}_{\mathrm{wpn}(T)})$ is a (P_2, w', i)-form with property (2) of (b). □

Now we complete the proof of the exact structure for all trees $T \in \mathcal{T}^{\mathrm{pl}}$.

Proof of Theorem 1.1.11 and Theorem 1.1.9 for all trees $T \in \mathcal{T}^{\mathrm{pl}}$. From Observation 5.2.11, to show that the number of bad graphs is much smaller than the number of good graphs it is enough to show that the number of Π-bad graphs is much smaller than the number of Π-good graphs. From Lemmas 5.2.38, 5.2.39, 5.2.40, 5.2.41, Observation 5.2.42, Corollary 5.2.46, and Lemmas 5.2.47, 5.2.48 we have the following. Almost all Π-bad graphs G have a set $W(G) \subset V(G)$ such that $|W(G)| \leq 2\,\mathrm{wpn}(T) \cdot (\log n)^2$, for there are indices $i_1 \neq i_2$ such that $G[\pi_{i_1} \setminus W(G)] \in \mathcal{G}_1$ and $G[\pi_{i_2} \setminus W(G)] \in \mathcal{G}_1$, and for $i \in \{i_1, i_2\}$, $\overline{G}[\pi_i \setminus W(G)]$ contains at most γn singleton vertices for any fixed $\gamma \in \left(0, \frac{1}{2\,\mathrm{wpn}(T)}\right)$, moreover, for all $j \in [\mathrm{wpn}(T)] \setminus \{i_1, i_2\}$, $G[\pi_j \setminus W(G)]$ is a clique. Let $\mathscr{D}$ be this collection of graphs. Let $\gamma \in \left(0, \frac{1}{2^{4t^2}\,\mathrm{wpn}(T)^3}\right)$ and let $\mathcal{F}' = \mathcal{F}(\mathscr{D})$, then by Lemma 5.2.49, $\mathcal{F}'$ is (τ, λ, k)-ordinary $P(T)$-free subsequence for $\tau = \frac{1}{w^2}$, any $\lambda > 0$ small enough and any $k \in \mathbb{N}$. Finally, by Lemma 3.3.14, the number of the remaining Π-bad graphs is much smaller than the number of Π-good graphs. □

Trees in $\mathcal{T}^{\text{npl}}$

We showed in Theorems 1.1.12, 1.1.14 and 1.1.16 that for any $T \in T^{\text{npl}}$ so $T \neq P_6$, the families $\mathcal{F}_i$, in a $\mathcal{P}(T)$-free sequence are either of the families $\mathcal{G}_\iota$, $\iota \in \{4, 5\}$ or the set of all cliques. In the case that $T = P_6$, then the families are either $\mathcal{G}_1$, $\mathcal{G}_6$, the sets of all stable set or the set of all cliques. We recall the definitions of the families $\mathcal{G}_\iota$, $\iota \in \{4, 5, 6\}$.

The family $\mathcal{G}_4 = \text{Forb}(\{P_4, 2K_2, K_2 + S_2, P_3 + S_1\})$ is the family of all graphs which are joins of graphs which are either a stable set or a disjoint union of a vertex and a clique. The complement of a graph in $\mathcal{G}_4$ is a disjoint union of stars and cliques.

The family $\mathcal{G}_5 = \text{Forb}(\{P_4, 2K_2, K_2 + S_2, P_3 + S_1, S_4\})$ is the family of all graphs which are joins of graphs which are either a stable set of size 3 or a disjoint union of a vertex and a clique (note that stable sets of size 1 and 2 are disjoint union of a vertex and a clique). The complement of a graph in $\mathcal{G}_5$ is the disjoint union of stars and cliques of size 3. We refer to a clique of size exactly 3 as triangle. Note that $\mathcal{G}_5 \subset \mathcal{G}_4$.

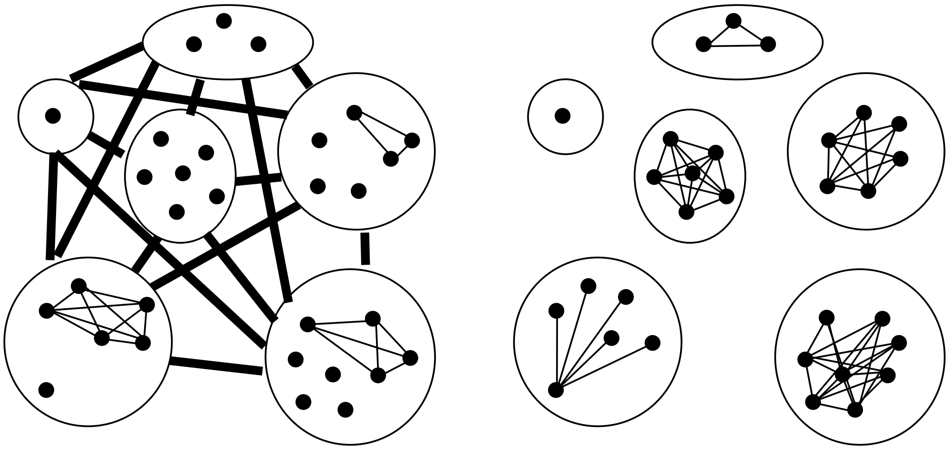
The family $\mathcal{G}_6 = \text{Forb}(\{P_4, 2K_2, P_3 + S_1\})$ is the family of graphs which are joins of graphs which are disjoint union of a clique and a stable set. Note that $\mathcal{G}_5 \subset \mathcal{G}_4 \subseteq \mathcal{G}_6$. See in Figure 5.11 an example for a graph G in $\mathcal{G}_6$ and its complement. In the figure, a thick edge between two sets of vertices represents the existence of all the possible edges between those two sets.

As before, we first give bounds on the number of graphs in the above families.

Lemma 5.2.50. *Let $n \in \mathbb{N}$ and $\iota \in \{4, 5, 6\}$,*

$$\text{Bell}(n) \leq |(\mathcal{G}_\iota)_n| \leq n! \cdot 2^{2n}.$$

Proof. Let $k \in [n]$ and let $(\pi_1, \pi_2, ..., \pi_k)$ be a partition of $[n]$ into k parts. It is possible to define a graph $\overline{G}$ for $G \in \mathcal{G}_5$ with respect to this partition by building a star from every part π_i, $i \in [k]$ (choosing the centre of every star arbitrary). In this way for two different partitions of $[n]$ we obtain two different graphs in $\mathcal{G}_5$. Therefore $|(\mathcal{G}_6)_n| \geq |(\mathcal{G}_4)_n| \geq |(\mathcal{G}_5)_n| \geq \text{Bell}(n)$.



(a) $G \in \mathcal{G}_6$ (b) The complement $\overline{G}$

Figure 5.11: A graph G in $\mathcal{G}_6$ and its complement .

It is possible to obtain any graph in $\mathcal{G}_6$ by considering some permutation of n and a choice of two subsets of n, let I_1, I_2 be those subsets. The elements of I_1 mark the partition into the different components of the complement graph. More precisely, let $i, i' \in I_1$ be such that i appears before i' and there is no other element of I_1 between i and i', then i together with all the elements between i and i' after i in the partitions define a component in the complement graph. An element I_2 which appears between $i, i' \in I_1$ marks the partition into a stable set and a clique. Let $j \in I_2$, then the elements between i and j are the elements of the stable set in the complement. Similarly, the elements between j to i' are the elements of the clique. $\square$

Let $G \in \mathcal{G}_6$, we call a connected component in $\overline{G}$ a **rich** component if it is of size at least 2. Note that a rich component of size at least 3 contains either an induced $\overline{P_3}$ or S_3. Let $r = r(G)$ be the number of rich components in G. Let $(\mathcal{G}_6)_n^{r,n',n''} \subseteq (\mathcal{G}_6)_n$ be the set of graphs G such that $\overline{G}$ has the following properties.

- $\overline{G}$ contains at most r rich components,
- the number of vertices in the rich components in $\overline{G}$ is at most n',
- there is a rich component with at least n'' vertices,

Similarly to before we denote $(\overline{\mathcal{G}_6})_n^{r,n',n''} = \{\overline{G} \mid G \in (\mathcal{G}_6)_n^{r,n',n''}\}$. If there is no restriction on any of the values in the upper script, then we write a $*$ is the corresponding entry.

Lemma 5.2.51. *Let $r, n', n'', n \in \mathbb{N}$, then*

$$|(\mathcal{G}_6)_n^{r,*,*}| \leq 2^{2n} \cdot r^n. \tag{5.4}$$

$$|(\mathcal{G}_6)_n^{r,n',*}| \leq 2^{2n} \cdot r^{n'}. \tag{5.5}$$

$$|(\mathcal{G}_6)_n^{*,*,n''}| \leq 2^{2n} \cdot n'' \cdot (n-n'')! \cdot 2^{2(n-n'')}. \tag{5.6}$$

Proof. Firstly we consider $(\mathcal{G}_6)_n^{r,*,*}$. There are at most 2^n ways to choose the vertices which are in the rich components. There are at most 2^n ways to choose the vertices which are in a stable set in their corresponding component in the complement. There are at most r^n ways to partition the elements chosen for the rich components into the different r rich components. Therefore we get the bound

$$|(\mathcal{G}_6)_n^{r,*,*}| \leq 2^{2n} \cdot r^n.$$

Secondly we consider $(\mathcal{G}_6)_n^{r,n',*}$. Similarly to before, there are at most 2^n ways to choose the n' vertices which are in the rich components. There are at most 2^n ways to choose the vertices which are in a stable set in their corresponding component in the complement. There are at most $r^{n'}$ ways to partition the elements chosen for the rich components into the different r rich components. Therefore we get the bound

$$|(\mathcal{G}_6)_n^{r,n',*}| \leq 2^{2n} \cdot r^{n'}.$$

Lastly we consider $(\mathcal{G}_6)_n^{*,*,n''}$. Let C be the rich component with at least n'' vertices. There are at most 2^n ways to choose the vertices in C, there are at most $2^{n''}$ ways to vertices in the stable set in C. By the bound in Lemma 5.2.50, the number of ways to choose a graph from $\mathcal{G}_6$ on the rest $(n-n'')$ vertices is at most $\cdot(n-n'')! \cdot 2^{2(n-n'')}$. Therefore we get the bound

$$|(\mathcal{G}_6)_n^{*,*,n''}| \leq 2^{2n} \cdot n'' \cdot (n-n'')! \cdot 2^{2(n-n'')}.$$

□

Lemma 5.2.52. *Let $n \in \mathbb{N}$ and $s, k \in [n]$, let $(\mathcal{G}_6^{s,k})_n \subset (\mathcal{G}_6)_n$ be the collection of graphs G which contain more than s rich components in $\overline{G}$ and the number of vertices in the union of all rich components, but the s largest, is k. Then,*

$$|(\mathcal{G}_6^{s,k})_n| \leq 2^{2n} \cdot s^n \cdot k^k.$$

Proof. There are at most 2^n ways to choose the vertices for the s largest rich components. There are at most s^n ways to partition the vertices into s components. There are at most k^k ways to partition the k remaining vertices into different components. There are at most 2^n ways to choose the vertices which are part of a stable set in their corresponding component. □

Lemma 5.2.53. *Let $n \in \mathbb{N}$, the number of graphs G in $(\mathcal{G}_6)_n$ so $\overline{G}$ contains at most $(\log n)^2$ rich components of size at least 3 is at most*

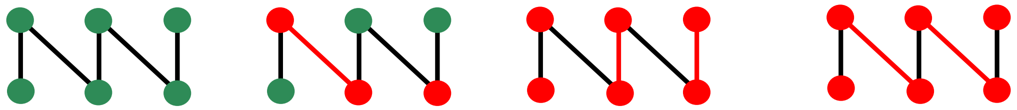
$$2^{2n} \cdot n^{\frac{n}{2}} \cdot (\log n)^{2n}.$$

Proof. There are at most 2^n ways to choose the vertices which are in components of size at most 2. By Theorem 5.2.22, there are at most $n^{\frac{n}{2}}$ ways to partition those vetices into components of size at most 2. There is at most $(\log n)^{2n}$ ways to assign vertices to the rich components of size at least 3. There are at most 2^n ways to choose which vertices are a part of a stable set in their corresponding component. □

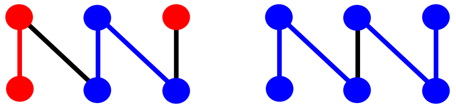
Lemma 5.2.54. *Let $n \in \mathbb{N}$, $\iota \in \{4, 5, 6\}$, and $n', c \in [n]$ then the number of graphs in $(\mathcal{G}_\iota)_n$ such that there are at least n' vertices in components of size at least c is at most*

$$2^{2n} \cdot \left(\frac{n}{c}\right)^{n'} \cdot |(\mathcal{G}_\iota)_{n-n''}|.$$

Proof. There are at most 2^n ways to choose the n' vertices which are in the components of size at least c. There are at most $\frac{n}{c}$ different components of size at least c and there are at most $\left(\frac{n}{c}\right)^{n'}$ ways to partition the n' vertices into those components. There are at most 2^n ways to choose the vertices which are a part of a stable set in each of the components in the complement. The graph on the rest $n - n'$ vertices is taken from $(\mathcal{G}_\iota)_{n-n''}$. □



(a) S_3 and S_3 (b) $\overline{P_3}$ and S_3 (c) $\overline{P_3}$ and $\overline{P_3}$



(d) $\overline{P_3}$ and P_3 (e) P_3 and P_3

Observation 5.2.55. *Let* $T \in \mathcal{T}^{\mathrm{npl}}$ *and* $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(T)})$ *be a partition of* $[n]$*. Let* $n_1 = |\pi_1|$ *then,*

$$|F(T, \Pi)| \geq |(\mathcal{G}_5)_{n_1}|.$$

The proof of the typical structure for trees in $\mathcal{T}^{\mathrm{npl}}$ is very similar to the proof of the typical structure of trees in $\mathcal{T}^{\mathrm{pl}}$.

Observation 5.2.56. *All the following pairs of graphs can be extended to* P_6 *with a proper choice of edges between the graphs in the pair,*

(a) (S_3, S_3) *with an edge fixed between them or without such edge, see Figure 5.12a.*

(b) $(\overline{P_3}, S_3)$ *with either an edge between the centre of* $\overline{P_3}$ *to one of vertices of the* S_3 *or without such edge, see Figure 5.12b.*

(c) $(\overline{P_3}, \overline{P_3})$ *with either an edge or without an edge between their centres, see Figure 5.12c.*

(d) $(\overline{P_3}, P_3)$ *with either an edge between the centre of the* $\overline{P_3}$ *to one of the ends of the* P_3 *or without such edge, see Figure 5.12d.*

(e) (P_3, P_3) *with one end of one of the* P_3 *not adjacent to any of the vertices of the other* P_3*, or adjacent to exactly one end of the other* P_3*, see Figure 5.12e.*

Using parts (a),(b),(c) and (d) of observation 5.2.56, and similarly to the proof of Lemma 5.2.31, it is possible to show the following.

Lemma 5.2.57. *Let K_1 and K_2 be such that $\overline{K_1}$ and $\overline{K_2}$ are rich components on k_1 and k_2 vertices respectively and $k_1 \geq 3, k_2 \geq 3$. The number of ways to choose edges between $V(K_1)$ and $V(K_2)$ without creating an induced copy of P_6 is at most*

$$2^{k_1 \cdot k_2} \cdot 2^{-\min\left\{\frac{k_1-1}{2^6}, \frac{k_2-1}{2^6}\right\}}.$$

Lemma 5.2.58. *Let $G_1, G_2 \in \mathcal{G}_6$ be two disjoint graphs. Let $|V(G_1)| = n_1$, $|V(G_2)| = n_2$ and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be maximum collections of disjoint rich components in $\overline{G}_1$ and $\overline{G}_2$, respectively. Let $s_1 = |\mathcal{S}_1|, s_2 = |\mathcal{S}_2|$ and $S_1 = \cup_{S \in \mathcal{S}_1} S, S_2 = \cup_{S \in \mathcal{S}_2} S$. We make the following assumptions.*

(i) No rich component in $\mathcal{S}_2$ contains at least $n_2 - s_2$ vertices, and

(ii) every rich component in $\mathcal{S}_2$ is larger than any rich component in $\mathcal{S}_1$.

The number of ways to choose edges between $V(G_1)$ and $V(G_2)$ without creating an induced copy of P_6 is at most $2^{n_1 \cdot n_2} \cdot 2^{-c|S_1| \cdot s_2}$ for some constant $c > 0$.

Proof. Let $\mathcal{S}'_1 \subseteq \mathcal{S}_1$ and $\mathcal{S}'_2 \subseteq \mathcal{S}_2$ be the collections of all rich components of size at least 3. Let $K_1 \in \mathcal{S}_1$ and $K_2 \in \mathcal{S}_2$. Then using Lemma 5.2.57, the number of ways to choose edges between $V(K_1)$ and $V(K_2)$ without creating P_6 is at most $2^{|V(K_1)| \cdot |V(K_2)|} \cdot 2^{-c|V(K_1)|}$. If $|\cup_{S \in \mathcal{S}'_1} S| \geq \frac{S_1}{2}$ then $|\mathcal{S}'_2| = s_2$ and we get the required bound. Hence we analyze the situation where this is not the case.

In the case that $|\cup_{S \in \mathcal{S}'_1} S| \leq \frac{S_1}{2}$, there are at least $\frac{S_1}{2}$ vertices in components of size exactly 2, let $\mathcal{T}_1$ be the maximal collection of disjoint P_3 in G_1. Using assumption (i), we can find in G_2 at least $\frac{s_2}{2}$ disjoint copies of $\overline{P_3}$ and P_3. Indeed we can define a set $\mathcal{T}_2$ by taking a copy of $\overline{P_3}$ from every component in $\mathcal{S}_2$ which contains $\overline{P_3}$ in G_2. Moreover, for every two sets in $\mathcal{S}_2$ which are stable sets in G_2, we can define a P_3 by taking a non-edge from one of such stable sets and a vertex from the other. Because, by observation 5.2.8, P_6 can be partitioned both into $(P_3, \overline{P_3})$ and (P_3, P_3), then we need to forbid at least one edge arrangement between every set $T_1 \in \mathcal{T}_1$ and $T_2 \in \mathcal{T}_2$. Therefore the number of ways to choose edges between $V(G_1)$ and $V(G_2)$ without creating an induced copy of P_6 is again at most $2^{n_1 \cdot n_2} \cdot 2^{-c|S_1| \cdot s_2}$ for some constant $c > 0$. □

We recall the structural properties of trees in $\mathcal{T}^{\mathrm{npl}}$ which are needed in the proof of the exact structure of almost all T-free graphs.

(A) Each of the graphs in $\{P_4, 2K_2, P_3 + K_1\}$ is $\frac{1}{4}$-universally extendable for all $T \in \mathcal{T}^{\mathrm{npl}}$. Moreover, $K_2 + S_2$ is $\frac{1}{4}$-universally extendable for all trees $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$ and S_4 is $\frac{1}{4}$-universally extendable for all trees $T \in \mathcal{T}^{\mathrm{npl}} \setminus \mathcal{S}$.

(B) P_6 is $(2, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\mathrm{npl}}$.

(C) Either P_3^* is $(2, \frac{1}{4})$-universally extendable or P_8 is $(3, \frac{1}{4})$-universally extendable for all $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$.

Those properties are shown in Lemmas 5.2.16, 5.2.17 and 5.2.18. As we did in the case of trees in $\mathcal{T}^{\mathrm{pl}}$, using those properties we give a few important corollaries to the general claims which we gave in Section 3.3.

Let $K \in \mathbb{N}$ and $\varepsilon > 0$ be the constants from Theorem 3.2.1 applied with $\mathrm{Forb}(T)$ and $\delta > 0$ sufficiently small. Let $\varepsilon' > 0$ be the constant from Theorem 3.2.3 applied with K. Let $\lambda \in [0, \frac{\min\{\varepsilon,\varepsilon'\}}{8})$.

Let $n \in \mathbb{N}$ be large enough and let $\rho > 0$ be the constant which we get from Theorem 3.2.2 for this n and $\xi > 0$ such that $2\xi \log \frac{1}{\xi} < \log e \cdot \frac{\lambda}{2^3 \cdot 2^{t^2+3} \cdot t \cdot \mathrm{wpn}(T)}$ where $t = V(T)$. Let $\Pi = (\pi_1, \pi_2, ..., \pi_{\mathrm{wpn}(T)})$ be a $\rho/4$-almost equal partition of $[n]$. We fix the partition Π for all of the following discussion. Let $\boldsymbol{n_i}= |\pi_i|$, $i \in [\mathrm{wpn}(T)]$.

Let G be a Π-conformal graph. Let $\mathcal{Y}(\Pi, G, i) = \mathcal{Y}(\Pi, G, \frac{1}{4}, i, \frac{1}{2^{t^2+1}})$ be the collection of sets obtained by adding greedily sets S which are $(\Pi, \frac{1}{4}, i, \frac{1}{2^{t^2+1}})$-linearly extremal. Let $\mathcal{Y}(G) = \cup_{i=1}^{\mathrm{wpn}(T)} \mathcal{Y}(\Pi, G, i)$, $Y = Y(G) = \cup_{Y \in \mathcal{Y}(G)} Y$ and $y = y(G) = |Y(G)|$ as defined in Section 3.3.

Corollary 5.2.59 (to Lemma 3.3.9)**.** *Let $T \in \mathcal{T}^{\mathrm{npl}}$, the number of Π-conformal graphs G, such that for some $i \in [\mathrm{wpn}(T)]$, $G[\pi_i \setminus Y(G)]$ contains an induced graph isomorphic to H' where $H' \in \{P_4, 2K_2, P_3 + S_1\}$ is much smaller than the number of Π-good graphs.*

Also, let $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$*, the number of* Π*-conformal graphs* G*, such that for some* $i \in [\mathrm{wpn}(T)]$*,* $G[\pi_i \setminus Y(G)]$ *contains an induced graph isomorphic to* $K_2 + S_2$ *is much smaller than the number of* Π*-good graphs.*

Moreover, if $T \in \mathcal{T}^{\mathrm{npl}} \setminus \mathcal{S}$*, then the number of* Π*-conformal graphs such that additionally there is an index* $i \in [\mathrm{wpn}(T)]$ *where* $G[\pi_i \setminus Y(G)]$ *contains an induced graph isomorphic to* S_4 *is much smaller than the number of* Π*-good graphs.*

Corollary 5.2.60 (to Corollary 3.3.10 and Theorem 2.2.4)**.** *Let* $T \in \mathcal{T}^{\mathrm{npl}}$*, then there is a constant* $C(T) > 0$ *which depends only on* T *such that the number of* Π*-conformal graphs where* $y \geq C(T) \log n$ *is much smaller than the number of* Π*-good graphs.*

Corollary 5.2.61 (to Lemma 3.3.12)**.** *Let* $T \in \mathcal{T}^{\mathrm{npl}}$*, the number of* Π*-conformal graphs where there are indices* $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$*, and there are subgraphs* J_{i_1} *and* J_{i_2} *in* $G[\pi_{i_1} \setminus Y(G)]$ *and* $G[\pi_{i_2} \setminus Y(G)]$*, respectively, such that* $G[V(J_{i_1}) \cup V(J_{i_2})]$ *is isomorphic to* P_6 *is much smaller than the number of* Π*-good graphs.*

Corollary 5.2.62 (to Lemma 3.3.13)**.** *Let* $T \in \mathcal{T}^{\mathrm{npl}}$*, and let* (J_1, J_2) *be a partition of* P_6*. The number of* Π*-conformal graphs where there indices* $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$*, such that* $G[\pi_{i_1} \setminus Y(G)]$ *contains at least* $f_1(n)$ *disjoint copies of a graph* J_1 *and* $G[\pi_{i_2} \setminus G(Y)]$ *contains at least* $f_2(n)$ *disjoint copies of a graph* J_2 *such that* $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$*, is much smaller than the number of* Π*-good graphs.*

Moreover, if $T \in \mathcal{T}^{\mathrm{npl}}$ *and* T *is not a path then the following holds. Let* (J_1, J_2) *be a partition of* P_3^**. The number of* Π*-conformal graphs where there indices* $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$*, such that* $G[\pi_{i_1} \setminus Y(G)]$ *contains at least* $f_1(n)$ *disjoint copies of a graph* J_1 *and* $G[\pi_{i_2} \setminus G(Y)]$ *contains at least* $f_2(n)$ *disjoint copies of a graph* J_2 *such that* $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$*, is much smaller than the number of* Π*-good graphs.*

Corollary 5.2.63 (to Lemma 3.3.13)**.** *Let* T *be a path of lengths at least* 8*, and let* (J_1, J_2, J_3) *be a partition of* P_8*. The number of* Π*-conformal graphs where there indices* i_1, i_2, i_3 *in* $[\mathrm{wpn}(T)]$*, such that* $G[\pi_{i_1} \setminus G(Y)]$ *and* $G[\pi_{i_2} \setminus G(Y)]$ *contain at least* $f_1(n)$ *disjoint copies of a graph* J_1 *and* J_2 *respectively, and* $G[\pi_{i_3} \setminus G(Y)]$ *contains at least* $f_2(n)$ *disjoint copies of a* *graph J_3 such that $f_1(n) \geq f_2(n)$ and $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$, is much smaller than the number of* Π-*good graphs.*

Let $\mathscr{C}$ be the set of all Π-conformal graphs and let $\mathscr{G}$ be the set of all Π-good graphs. Similarly to before, let $\mathscr{C}' \subset \mathscr{C}$ be the set of all Π-conformal graphs G with the following properties.

(a) For all $T \in \mathcal{T}^{\mathrm{npl}}$, $G[\pi_i \setminus Y(G)] \in \mathcal{G}_6$, $i \in [\mathrm{wpn}(T)]$, for all $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$, $G[\pi_i \setminus Y(G)] \in \mathcal{G}_4$, $i \in [\mathrm{wpn}(T)]$, and for all $T \in \mathcal{T}^{\mathrm{npl}} \setminus \mathcal{S}$, $G[\pi_i \setminus Y(G)] \in \mathcal{G}_5$, $i \in [\mathrm{wpn}(T)]$.

(b) $y \leq C(T) \log n$ for some constant $C(T) > 0$.

(c) There are no indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, where there are subgraphs J_{i_1} and J_{i_2} in $G[\pi_{i_1} \setminus Y(G)]$ and $G[\pi_{i_2} \setminus Y(G)]$, respectively, such that $G[J_{i_1} \cup J_{i_2}]$ is isomorphic to P_6. Moreover, if T is not a path then the above is also true for the case where $G[J_{i_1} \cup J_{i_2}]$ is isomorphic to P_3^*.

(d) There are no indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, such that $G[\pi_{i_1} \setminus Y(G)]$ contains at least $f_1(n)$ disjoint copies of a graph J_1 and $G[\pi_{i_2} \setminus Y(G)]$ contains at least $f_2(n)$ disjoint copies of a graph J_2 such that the following holds. The collection (J_1, J_2) is a partition of P_6 and, $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$. Moreover, if T is not a path then the above is also true for the case where (J_1, J_2) is a partition of P_3^*.

(e) For a tree T which is path of length at least 8, there are no indices $i_1 \neq i_2 \neq i_3 \in [\mathrm{wpn}(T)]$, such that $G[\pi_{i_1} \setminus Y(G)]$ and $G[\pi_{i_2} \setminus Y(G)]$ contain at least $f_1(n)$ disjoint copies of a graph J_1 and J_2 respectively, and $G[\pi_{i_3} \setminus Y(G)]$ contains at least $f_2(n)$ disjoint copies of a graph J_3 such that the following holds. The collection (J_1, J_2, J_3) is a partition of P_8, $f_1(n) \geq f_2(n)$ and $\frac{n \log n}{f_1(n) \cdot f_2(n)} = \mathrm{o}(1)$.

Lemma 5.2.64. *The number of graphs in $\mathscr{C} \setminus \mathscr{C}'$ is much smaller than the number of graphs in $\mathscr{G}$.*

Proof. This is a direct corollary to Corollaries 5.2.59, 5.2.60, 5.2.61, 5.2.62 and 5.2.63. ☐

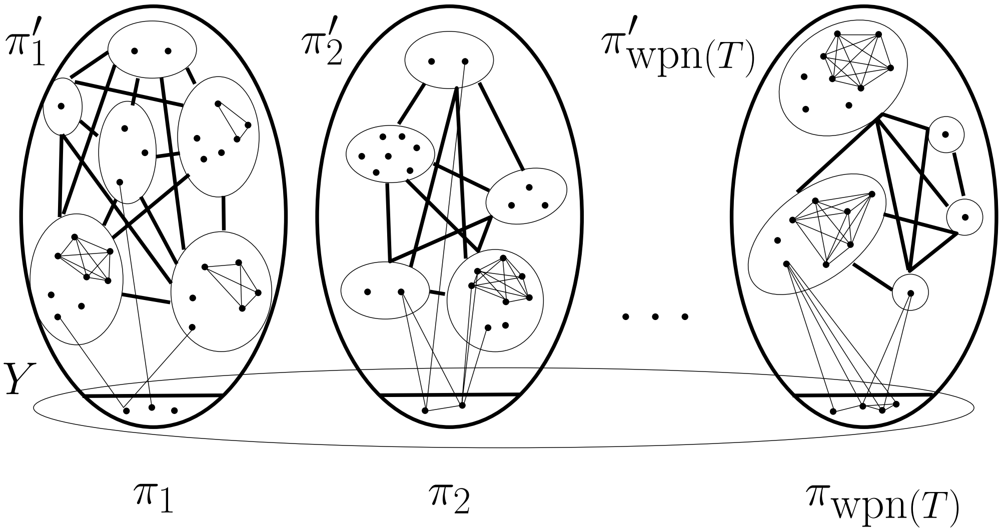


Figure 5.13: A sketch of a graph $G \in \mathscr{B}$. The edges between the parts are not drawn.

Let $G \in \mathscr{C}'$ then by the definition of the set $\mathscr{C}'$, $G[\pi_i \setminus Y(G)] \in \mathcal{G}_6$, $i \in [\mathrm{wpn}(T)]$, therefore $\overline{G}$ is a disjoint union of rich components and singleton vertices. Hence for each $i \in [\mathrm{wpn}(T)]$, we can define $r_i = r_i(\overline{G}[\pi_i \setminus Y(G)])$ to be the number of rich components in $\overline{G}[\pi_i \setminus Y(G)]$ and similarly we define $\ell_i = \ell_i(\overline{G}[\pi_i \setminus Y(G)])$ to be the maximum number of connected components which are edges in $\overline{G}[\pi_i \setminus Y(G)]$.

We start from the Π-bad graphs in the set $\mathscr{C}'$, let $\mathscr{B} \subseteq \mathscr{C}'$ be the collection of all Π-bad graphs in $\mathscr{C}'$. Note that we will use the fact that we are considering Π-bad graphs only towards the end of the proof. See Figure 5.13 for a sketch of an example of a graph in $\mathscr{B}$. In particular, in Figure 5.13, $r_1 = 5$, $r_2 = 5$, $r_{\mathrm{wpn}(T)} = 2$. For $G \in \mathscr{B}$, for each $i \in [\mathrm{wpn}(T)]$, $\pi'_i = \pi_i \setminus Y(G)$.

Let $\mathscr{F}$ be a family of graphs G which are defined with respect to the properties of the subgraphs $G[\pi_i]$, $i \in [\mathrm{wpn}(T)]$. As before we define $\mathcal{F}(\mathscr{F}) = (\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ to be the collection of families where $\mathcal{F}_i = \cup_{G \in \mathscr{F}} G[\pi_i]$, $i \in [\mathrm{wpn}(T)]$.

Let $\mathscr{B}_{\mathbf{1}} \subset \mathscr{B}$ be the set of graphs in $\mathscr{B}$ such that there is no index $i \in [\mathrm{wpn}(T)]$ where $r_i \geq \frac{n}{(\log n)^2}$.

Lemma 5.2.65. *Let $T \in \mathcal{T}^{\mathrm{npl}}$, the number of graphs in $\mathscr{B}_1$ is much smaller than the number of Π-good graphs $\mathscr{G}$.*

Proof. We start by listing the different subsets of graphs in $\mathscr{B}_1$ that we consider during the proof. We show that the number of graphs in each of the following subsets is much smaller than the number of graphs in $\mathscr{G}$. As before, we finish by showing the number of graphs in $\mathscr{B}_1$ which are not in any of the following families is also much smaller the number of graphs in $\mathscr{G}$.

Let $\beta \in \left(0, \frac{1}{2^8 \operatorname{wpn}(T)^2}\right)$. Let $\mathcal{B}_1 = \mathcal{B}_1(\beta) \subset \mathscr{B}_1$ be the set of graphs G which have at most n^β rich components (components of size at least 2) in $\overline{G}[\pi'_1] \cup \overline{G}[\pi'_2] \cup ... \cup \overline{G}[\pi'_{\operatorname{wpn}(T)}]$.

Let $\mathcal{S}$ be the set of the n^β largest rich components in $\overline{G}[\pi'_1] \cup \overline{G}[\pi'_2] \cup ... \cup \overline{G}[\pi'_{\operatorname{wpn}(T)}]$. Let S be the union of the vertices in the components in $\mathcal{S}$. Let $\mathcal{B}_2 \subset \mathscr{B}_1 \setminus \mathcal{B}_1$ be the set of Π-bad graphs where each subgraph $\overline{G}[\pi'_i \setminus S]$, $i \in [\operatorname{wpn}(T)]$, has at least one of the following properties,

(1) it contains at most n^β rich components,

(2) there are at most βn_i vertices in rich components,

(3) there are at least $(1-\beta)n_i$ vertices in one component.

Next we consider the graphs in $\mathscr{B}_1 \setminus \cup_{i=1}^2 \mathcal{B}_i$. By the definition of the graphs in this set there must be an index $i \in [\operatorname{wpn}(T)]$, such that $\overline{G}[\pi'_i \setminus S]$ does not have any of the properties (1),(2) and (3) as above, without loss of generality $i = 1$. Let $\delta \in (0,1)$ and let $\mathcal{B}_3 \subset \mathscr{B}_1 \setminus \cup_{i=1}^2 \mathcal{B}_i$ be the set of graphs so there is an index $i \geq 2$ where $G[\pi'_i]$ contains a stable set of size at least $n_i - n^\delta$.

Let $\mathcal{B}_4 \subset \mathscr{B}_1 \setminus \cup_{i=1}^3 \mathcal{B}_i$ be the family of graphs where for some $i \geq 2$, $\overline{G}[\pi'_i]$ contains more than $(\log n)^2$ components from $\mathcal{S}$.

We start by bounding the number of graphs in $\mathcal{B}_1 \subseteq \mathscr{B}_1$. Let $\mathcal{F}(\mathcal{B}_1) = (\mathcal{F}_1^1, \mathcal{F}_2^1, ..., \mathcal{F}_{\operatorname{wpn}(T)}^1)$. By the definition of $\mathcal{B}_1$, for each $i \in [\operatorname{wpn}(T)]$, $r_i \leq n^\beta$, so by Lemma 5.2.51, for each $i \in [\operatorname{wpn}(T)]$, $|\mathcal{F}_i^1| \leq |(\mathcal{G}_6)_{n_i}^{n^\beta,*,*}| \leq 2^{3n_i} \cdot n^{\beta n_i}$. Therefore by Lemma 3.3.8, the number of graphs in $\mathcal{B}_1$ is at most

$$\begin{aligned}
&2^m \cdot \left(\prod_{i=1}^{\mathrm{wpn}(T)} 2^{3n_i} \cdot n^{\beta n_i} \right) \cdot 2^{(b+3)n} \\
&\leq 2^m \cdot n^{\beta \,\mathrm{wpn}(T) n} \cdot 2^{(b+6)n} \\
&\leq 2^m \cdot n^{\frac{n}{2^7 \,\mathrm{wpn}(T)}}.
\end{aligned}$$

We compare it to the lower bound on the number of Π-good graphs from Observation 5.2.55.

$$\begin{aligned}
\frac{|\mathcal{B}_1|}{|\mathscr{G}|} &\leq \frac{2^m \cdot n^{\frac{n}{2^7 \,\mathrm{wpn}(T)}}}{2^m \cdot \mathrm{Bell}(n_1)} \leq \frac{n^{\frac{n}{2^7 \,\mathrm{wpn}(T)}}}{\left(\frac{\frac{n}{2 \,\mathrm{wpn}(T)}}{e \ln \frac{n}{2 \,\mathrm{wpn}(T)}} \right)^{\frac{n}{2 \,\mathrm{wpn}(T)}}} \\
&= \frac{\left(e \ln \frac{n}{2 \,\mathrm{wpn}(T)} \right)^{\frac{n}{2 \,\mathrm{wpn}(T)}}}{n^{\frac{(1-2^6)n}{2^7 \,\mathrm{wpn}(T)}}} = \mathrm{o}(1).
\end{aligned}$$

Next we consider the graphs in $\mathcal{B}_2 \subseteq \mathscr{B}_1 \backslash \mathcal{B}_1$. Let $\mathcal{F}(\mathcal{B}_2) = (\mathcal{F}_1^2, \mathcal{F}_2^2, ..., \mathcal{F}_{\mathrm{wpn}(T)}^2)$. By the definition of $\mathcal{B}_2$, for each $i \in [\mathrm{wpn}(T)]$, $|\mathcal{F}_i^2| \leq 2^{n_i} \cdot n^{\beta n_i} \cdot \max\{|(\mathcal{G}_6)_{n_i}^{n^\beta,*,*}|, |(\mathcal{G}_6)_{n_i}^{*,\beta n_i,*}|, |(\mathcal{G}_6)_{n_i}^{*,*,(1-\beta)n_i}|\}$. By Lemma 5.2.51, $|(\mathcal{G}_6)_{n_i}^{n^\beta,*,*}| \leq 2^{2n_i} \cdot n^{\beta n_i}$; $|(\mathcal{G}_6)_{n_i}^{*,\beta n_i,*}| \leq 2^{2n} \cdot n^{\beta n_i}$; $|(\mathcal{G}_6)_{n_i}^{*,*,(1-\beta)n_i}| \leq 2^{2n_i} \cdot n_i \cdot (\beta n_i)^{\beta n_i}$. Hence for all $i \in [\mathrm{wpn}(T)]$, $|\mathcal{F}_i^2| \leq 2^{5n_i} \cdot n^{2\beta n_i}$. By Lemma 3.3.8, the number of graphs in $\mathcal{B}_2$ is at most

$$\begin{aligned}
&2^m \cdot \left(\prod_{i=1}^{\mathrm{wpn}(T)} 2^{5n_i} \cdot n^{2\beta n_i} \right) \cdot 2^{(b+3)n} \\
&\leq 2^m \cdot n^{\beta \,\mathrm{wpn}(T) n} \cdot 2^{(b+8)n} \\
&\leq 2^m \cdot n^{\frac{n}{2^7 \,\mathrm{wpn}(T)}}.
\end{aligned}$$

As before, the above number is much smaller than the lower bound on the number of Π-good graphs from Observation 5.2.55.

Now we consider the graphs in $\mathcal{B}_3 \subseteq \mathscr{B}_1 \setminus \cup_{i=1}^2 \mathcal{B}_i$. Let $G \in \mathcal{B}_3$, then by the definition of this family there must be an index $i \in [\mathrm{wpn}(T)]$, such that $\overline{G}[\pi_i' \setminus S]$ contains at least n^β components of size at least 2, at least βn_i vertices are in components of size at least 2 and no component contains at least $(1-\beta)n_i$ vertices. Without loss of generality $i = 1$. Let $i \geq 2$

so $G[\pi'_i]$ contains a stable set of size at least $n_i - n^{\delta}$. This means that in particular $G[\pi'_i]$ contains at least $\frac{n}{8\,\mathrm{wpn}(T)}$ disjoint S_3.

If $T \neq P_6$, then by properties (d) and (e) of graphs in $\mathscr{B}_1$, we can deduce that the number of such graphs is much smaller than the number of Π-good graphs. Hence assume that $T = P_6$. By property (d) of graphs in $\mathscr{B}_1$, $G[\pi'_1]$ can contain at most $(\log n)^2$ disjoint copies of either S_3 or $\overline{P_3}$. Let $\mathcal{F}(\mathcal{B}_i) = (\mathcal{F}^3_1, \mathcal{F}^3_2)$. Then $|\mathcal{F}^3_1| \leq 2^{2n_1} \cdot n_1^{\frac{n_1}{2}} \cdot (\log n)^{2n_1}$ and by Lemma 5.2.51, $|\mathcal{F}^3_2| \leq 2^{4n_2} \cdot (n_i - n^{\delta}) \cdot n^{\delta n^{\delta}}$. By Lemma 3.3.8, the number of graphs in $\mathcal{B}_3$ in the case that $T = P_6$ is at most

$$\begin{aligned}
&2^m \cdot 2^{2n_1} \cdot n_1^{\frac{n_1}{2}} \cdot (\log n)^{2n_1} \cdot 2^{4n_2} \cdot (n_i - n^{\delta}) \cdot n^{\delta n^{\delta}} \cdot 2^{(b+3)n} \\
&2^m \cdot 2^{(b+11)n} \cdot n_1^{\frac{n_1}{2}} \cdot (\log n)^{2n_1} \cdot n^{\delta n^{\delta}}.
\end{aligned}$$

We compare it to the lower bound on the number of Π-good graphs from Observation 5.2.55.

$$\begin{aligned}
\frac{|\mathcal{B}_3|}{|\mathscr{G}|} &\leq \frac{2^m \cdot 2^{(b+11)n} \cdot n_1^{\frac{n_1}{2}} \cdot (\log n)^{2n_1} \cdot n^{\delta n^{\delta}}}{2^m \cdot \mathrm{Bell}(n_1)} \\
&\leq \frac{2^{(b+11)n} \cdot n_1^{\frac{n_1}{2}} \cdot 2^{3n \log\log n}}{\left(\frac{n_1}{e \ln n_1}\right)^{n_1}} \leq \frac{2^{10n \log\log n}}{\left(\frac{n}{2\,\mathrm{wpn}(T)}\right)^{\frac{n}{2\,\mathrm{wpn}(T)}}} = \mathrm{o}(1).
\end{aligned}$$

Next we consider the graphs in $\mathcal{B}_4 \subset \mathscr{B}_1 \setminus \cup_{i=1}^{3} \mathcal{B}_i$. Let $i \geq 2$ be such $\overline{G}[\pi'_i]$ contains more than $(\log n)^2$ components from $\mathcal{S}$. Note that as mentioned earlier $\mathscr{B}_1 \subset \mathscr{B}$, therefore any $G \in \mathcal{B}_4$ has property (c), that is, there are no indices $i_1 \neq i_2 \in [\mathrm{wpn}(T)]$, where there are subgraphs J_{i_1} and J_{i_2} in $G[\pi'_{i_1}]$ and $G[\pi'_{i_2}]$, respectively, such that $G[V(J_{i_1}) \cup V(J_{i_2})]$ is isomorphic to P_6. We apply Lemma 5.2.58 to the graphs induced on $G[\pi'_1]$ and $G[\pi'_i \cap S]$ and get that the number of ways to choose edges between those subgraphs is at most $2^{|\pi_1| \cdot |\pi_i \cap S|} \cdot 2^{-c\beta n \cdot (\log n)^2}$ for some constant $c > 0$. Let $\mathcal{F}(\mathcal{B}_4) = (\mathcal{F}^4_1, \mathcal{F}^4_2, ..., \mathcal{F}^4_{\mathrm{wpn}(T)})$, then because for each $\mathcal{F}^4_i \subseteq \mathrm{Forb}(P_4)$ and Theorem 2.2.4, we have that $|\mathcal{F}^4_i| \leq 2^{3n_i \log n_i}$. In total, by Lemma 3.3.8, the number of graphs in $\mathcal{B}_4$ is at most

$$\begin{aligned}
&2^m \cdot \left(\prod_{i=1}^{\mathrm{wpn}(T)} 2^{3n_i \log n_i} \right) \cdot 2^{(b+3)n} \cdot 2^{-c\beta n \cdot (\log n)^2} \\
&\leq 2^m \cdot 2^{3n \log n} \cdot 2^{(b+3)n} \cdot 2^{-c\beta n \cdot (\log n)^2}.
\end{aligned}$$

The values of $c > 0$ and $\beta > 0$ are constants, therefore for n large enough, the above is much smaller then 2^m and therefore much smaller than the number of Π-good graphs.

Let $G \in \mathscr{B}_1 \setminus \cup_{i=1}^4 \mathcal{B}_i$, by the definition of this family, $\overline{G}[\pi'_1 \setminus S]$ contains at least n^β components of size at least 2, at least βn_1 vertices are in components of size at least 2 and no component contains at least $(1-\beta)n_1$ vertices. For each $i \geq 2$, $G[\pi'_i]$ does not contain a stable set of size at least $n_i - n^\delta$ and more than $(\log n)^2$ components from $\mathcal{S}$. Therefore $\overline{G}[\pi'_1]$ also contains at least $\frac{n^\beta}{2}$ components from $\mathcal{S}$.

Let $\mathcal{F}(\mathscr{B}_1 \setminus \cup_{i=1}^4 \mathcal{B}_i) = (\mathcal{F}_1^5, \mathcal{F}_2^5, ..., \mathcal{F}_{\mathrm{wpn}(T)}^5)$. First we want to bound $|\mathcal{F}_1^5|$. By Lemma 5.2.51, $|\mathcal{F}_1^5| \leq 2^{n_1} \cdot \left(\frac{n}{(\log n)^2}\right)^{n_1}$. For each $i \geq 2$, let s_i be the number of components from $\mathcal{S}$ in $\overline{G}[\pi'_i]$ and let k_i be the number of vertices in rich components in $\overline{G}[\pi'_i \setminus S]$. Hemce by Lemma 5.2.52, for each $i \geq 2$, $|\mathcal{F}_i^5| \leq 2^{2n_i} n^{k_i} s_i^n$. We apply again Lemma 5.2.58 to the graphs induced on $G[\pi'_1]$ and $G[\pi'_i \cap S]$ and to the graphs induced on $G[\pi'_1 \cap \mathcal{S}]$ and $G[\pi'_i \setminus S]$ and get that the number of ways to choose edges between those subgraphs is at most $2^{|\pi_1 \setminus S| \cdot |\pi_i \cap S|} \cdot 2^{-c\beta n \cdot s_i} \cdot 2^{|\pi_1 \cap S| \cdot |\pi_i \setminus S|} \cdot 2^{-cn^\beta \cdot k_i}$ for each $i \geq 2$. Therefore by Lemma 3.3.8 we get the following bound on the number of graphs in $\mathscr{B}_1 \setminus \cup_{i=1}^4 \mathcal{B}_i$,

$$\begin{aligned} &2^m \cdot 2^{(b+3)n} \cdot n_1 \cdot 2^{n_1} \cdot \left(\frac{n}{(\log n)^2}\right)^{n_1} \cdot \prod_{i \geq 2} 2^{2n_i} n^{k_i} s_i^n \cdot 2^{-c\beta n \cdot s_i} \cdot 2^{-cn^\beta \cdot k_i} \\ &\leq 2^m \cdot 2^{(b+5)n} \cdot n_1 \cdot 2^{n_1} \cdot \left(\frac{n}{(\log n)^2}\right)^{n_1} \cdot 2^{(\log n - cn^\beta)k} \cdot 2^{(\mathrm{wpn}(T) \log n - c\beta n)s} \end{aligned}$$

where $k = \sum_{i \geq 2} k_i$ and $s = \sum_{i \geq 2} s_i$. We compare it to the lower bound of $2^m \cdot |(\mathcal{G}_5)_{n_1}|$ on the number of Π-good graphs, we also use the lower bound on $\mathcal{G}_5$ from Lemma 5.2.50. Therefore,

$$\begin{aligned} \frac{|\mathscr{B}_1 \setminus \cup_{i=1}^4 \mathcal{B}_i|}{|\mathscr{G}|} &\leq \frac{2^m \cdot 2^{(b+6)n} \cdot 2^{n_1} \cdot \left(\frac{n}{(\log n)^2}\right)^{n_1} \cdot 2^{(\log n - cn^\beta)k} \cdot 2^{(\mathrm{wpn}(T) \log n - c\beta n)s}}{2^m \cdot |(\mathcal{G}_5)_{n_1}|} \\ &\leq \frac{2^{(b+6)n} \cdot n^{n_1} \cdot 2^{(\log n - cn^\beta)k} \cdot 2^{(\mathrm{wpn}(T) \log n - c\beta n)s}}{(\log n)^{2n_1} \cdot \left(\frac{n}{2e\,\mathrm{wpn}(T) \log n_1}\right)^{n_1}} \\ &\leq \frac{2^{(b+6)n} \cdot \left(2e\,\mathrm{wpn}(T)\right)^{n_1} \cdot 2^{(\log n - cn^\beta)k} \cdot 2^{(\mathrm{wpn}(T) \log n - c\beta n)s}}{(\log n)^{n_1}} = \mathrm{o}(1). \end{aligned}$$

Hence the number of graphs in $\mathscr{B}_1$ is much smaller than the number of Π-good graphs.

□

Let $\mathscr{B}_{\mathbf{2}} \subseteq \mathscr{B} \setminus \mathscr{B}_1$ be the collection of graphs such that $r_1 \geq \frac{n}{(\log n)^2}$ and there is an index $i \geq 2$ so the number of vertices in the union of the stable sets in G in the rich components is at least $(\log n)^3$.

Lemma 5.2.66. *Let $T \in \mathcal{T}^{\mathrm{npl}}$, the number of graphs in $\mathscr{B}_2$ is much smaller than the number of* Π-*good graphs* $\mathscr{G}$*.*

Proof. As in the proof of Lemma 5.2.65, let $\delta \in (0,1)$, then no $G[\pi_i']$, $i \geq 2$, contains a stable set of size at least $n_i - n^\delta$.

Let $i \geq 2$ and let U_i be the union of all the vertices in the stable sets in $G[\pi_i']$. Then we can find in $G[\pi_i']$ at least $\min\{\frac{|U_i|}{2}, n^\delta\}$ disjoint copies of P_3. Then the number of such graphs is much smaller than the number of Π-good graphs by property (d) of graphs in $\mathscr{B}$. ☐

Let $G \in \mathscr{B} \setminus (\mathscr{B}_1 \cup \mathscr{B}_2)$. Let $W(G) \subseteq [n]$ be the minimal set such that there is an index $i \in [\mathrm{wpn}(T)]$, without loss of generality $i = 1$, such that for $T = P_6$, $G[\pi_1 \setminus W(G)] \in \mathcal{G}_6$, for all trees $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$, $G[\pi_1 \setminus W(G)] \in \mathcal{G}_4$, and for all trees $T \in \mathcal{T}^{\mathrm{npl}} \setminus (\mathcal{S} \cup \{P_6\})$, $G[\pi_1'] \in \mathcal{G}_5$. Moreover, for each $i \geq 2$, $G[\pi_i \setminus W(G)]$ is a clique.

Lemma 5.2.67. *Let $T \in \mathcal{T}^{\mathrm{npl}}$ and let $G \in \mathscr{B} \setminus \mathscr{B}_1$, then $|W(G)| \leq 2\,\mathrm{wpn}(T) \cdot (\log n)^3$.*

Proof. Let $T \in \mathcal{T}^{\mathrm{npl}}$, then for each $i \geq 2$, let U_i be the union of all the vertices in the stable sets in $G[\pi_i']$. We define $W(G) = \left(\cup_{i \geq 2} U_i\right) \cup Y$. By property (b) of the definition of $\mathscr{B}$, $y = |Y| \leq C(T) \log n$. By Lemma 5.2.66, for any $G \in \mathscr{B} \setminus \cup_{i=1}^{2} \mathscr{B}_i$, $|U_i| \leq (\log n)^3$ for all $i \geq 2$. Therefore $|W(G)| \leq 2\,\mathrm{wpn}(T) \cdot (\log n)^3$. ☐

Let $\mathscr{B}_{\mathbf{3}} \subseteq \mathscr{B} \setminus \cup_{i=1}^{2} \mathscr{B}_i$ be the set of graphs where $\overline{G}[\pi_1 \setminus W(G)]$ contains at least $\frac{n}{\log\log\log n}$ vertices in components of size at least $(\log n)^2$.

Lemma 5.2.68. *Let $T \in \mathcal{T}^{\mathrm{npl}}$, the number of graphs in $\mathscr{B}_3$ is much smaller than the number of* Π-*good graphs* $\mathscr{G}$*.*

Proof. Repeating the arguments in the proofs of Lemmas 5.2.65 and 5.2.66 we can obtain $\mathscr{B}_3' \subseteq \mathscr{B}_3$ such that the number of graphs in $\mathscr{B}_3 \setminus \mathscr{B}_3'$ is much smaller than the number of graphs in $\mathscr{G}$. Moreover, the following holds for every graphs $G \in \mathscr{B}'_3$. The subgraph $\overline{G}[\pi'_1]$ contains at least $\frac{n}{(\log n)^2}$ rich components, and the subgraphs $\overline{G}[\pi'_i]$, $i \geq 2$, contain at most $(\log n)^3$ vertices in the union of the stable sets in G in the rich components. Let $\iota \in \{4, 5, 6\}$, to count the graphs on $\pi_1 \setminus W(G)$ we use Lemma 5.2.54,

$$\frac{|\mathscr{B}_2|}{|\mathscr{G}|} \leq \frac{2^m \cdot 2^{(b+4)n} \cdot 2^{n_1} \cdot \left(\frac{n}{(\log n)^2}\right)^{\frac{n}{\log\log\log n}} \cdot |(\mathcal{G}_\iota)_{n_1 - \frac{n}{\log\log\log n}}| \cdot 2^{(\log n - cn^\beta)k} \cdot 2^{(\mathrm{wpn}(T)\log n - c\beta n)s}}{2^m \cdot \left(\frac{\frac{n}{\log\log\log n}}{e \log \frac{n}{\log\log\log n}}\right)^{\frac{n}{\log\log\log n}} |(\mathcal{G}_\iota)_{n_1 - \frac{n}{\log\log\log n}}|}$$

$$\leq \frac{2^{(b+6)n} \cdot 2^{(\log n - cn^\beta)k} \cdot 2^{(\mathrm{wpn}(T)\log n - c\beta n)s}}{(\log n)^{\frac{n}{\log\log\log n}}} = \mathrm{o}(1),$$

as required. $\square$

Let $\boldsymbol{\mathcal{F}' = (\mathcal{F}'_1, \mathcal{F}'_2, ..., \mathcal{F}'_{\mathrm{wpn}(T)})}$ be such $\mathcal{F}'_1 \subseteq \mathcal{G}_6$ and for any $G \in \mathcal{F}'_1$, the complement $\overline{G}$ contains at most $\frac{n}{\log\log\log n}$ vertices in components of size at least $(\log n)^2$. For every $i \geq 2$, $\mathcal{F}'_i$ is a collection of all cliques.

Lemma 5.2.69. *Let $\varepsilon, \varepsilon' > 0$ to be the constants defined earlier in the subsection. The sequence $\mathcal{F}'$ is (τ, λ, k)-ordinary $P(T)$-free subsequence for any $\lambda \in [0, \frac{\min\{\varepsilon, \varepsilon'\}}{8})$, $\tau = \frac{1}{w^2}$, and any $k \in \mathbb{N}$.*

Proof. We check properties (a) and (b) of the definition of (τ, λ, k)-ordinary $P(T)$-free subsequence. Let $Q = (G_1, G_2, ..., G_{\mathrm{wpn}(T)})$ be a sequence of graphs such that $G_i \in \mathcal{F}'_i$ for each $i \in [\mathrm{wpn}(T)]$.

First we check property (a). By the definition of $\mathcal{F}'_1$, $\overline{G}_1$ contains at most $\frac{n}{\log\log\log n}$ vertices in components of size at least $(\log n)^2$, then for $v \in V(G_1)$, $\deg_{\overline{G}_1}(v) = \overline{\deg}_{G_1}(v) \leq \frac{n}{\log\log\log n} \leq \tau n$ for any $\tau > 0$. Let $i \geq 2$, G_i is a cliques, therefore for $v \in V(G_i)$, $\overline{\deg}_{G_i}(v) = 0$. Let $\tau = \frac{1}{w^2}$.

Now we check property (b). Let $w' \notin \cup_{i \in [\mathrm{wpn}(T)]} V(G_i)$ be such for each $i \in [\mathrm{wpn}(T)]$, $G_i \cup \{w'\} \notin \mathcal{F}'_i$. Let $\ell_i(w') = \min\{|N(w') \cap V(G_i)|, |\overline{N}(w') \cap V(G_i)|\}$, $i \in [\mathrm{wpn}(T)]$, let $\ell(w') = \min_{i \in [\mathrm{wpn}(T)]} \ell_i(w')$ and let $i \in [\mathrm{wpn}(T)]$ be such that $\ell(w') = \ell_i(w')$, then for each $j \in [\mathrm{wpn}(T)] \setminus \{i\}$, $\ell_j(w') \geq \tau n$, that is w' is not (τ, G_j)-trivial.

First we consider the case where $i = 1$. Let $\mathcal{N}$ be either (i) a collection of disjoint copies of $\overline{P_3}$ or (ii) a collection of disjoint copies of P_3 in $\overline{N}(w') \cap V(G_1)$. Note that by the assumptions on G_1, $|\mathcal{N}| \geq \frac{\ell(w') - \frac{n}{\log\log\log n}}{2(\log n)^2}$ and for each $N \in \mathcal{N}$, $N \cup \{w'\}$ is isomorphic to $K_2 + S_2$ in case (i) and to $P_3 + K_1$ in case (ii). By our assumptions, for each $i \geq 2$, w' is not (τ, G_i)-trivial, therefore for each of the possible edge arrangements between a vertex and an edge, G_i contains a set $\mathcal{H}_i$ of at least $\frac{\tau n}{4}$ disjoint edges, so w' has this edge arrangement. Let $P = (H_1, H_2, ..., H_{\text{wpn}(T)})$ be a partition of T where $H_1 = K_2 + S_2$ in case (i) and $H_1 = P_3 + K_1$ in case (ii), for each $i \geq 2$, H_i is an edge. The tree T has such a partition by Lemma 5.2.2 and Observation 5.2.7. The collection of sets $(\mathcal{N}, \mathcal{H}_2, \mathcal{H}_3, ..., \mathcal{H}_{\text{wpn}(T)})$ is a (P, w', i)-form in Q with property (1) of (b).

Now, let $i \geq 2$. From Lemma 5.2.21, G_1 contains a set $\mathcal{P}$ of disjoint copies of P_3 such that either (i) w' is adjacent to exactly one end (and not other vertices) of each copy of P_3 in $\mathcal{P}$, or (ii) w' is not adjacent to any of the vertices of each copy of P_3 in $\mathcal{P}$. Moreover, $|\mathcal{P}| \geq \frac{\tau n - \frac{n}{\log\log\log n}}{2(\log n)^2}$. As before, by our assumptions, for each $j \in [\text{wpn}(T)] \setminus \{i\}$, w' is not (τ, G_j)-trivial, therefore for each of the possible edge arrangements between a vertex and an edge, G_j contains a set $\mathcal{H}_j$ of at least $\frac{\tau n}{4}$ disjoint edges, so w' has this edge arrangement.

Let $T \in \mathcal{T}^{\text{npl}}$, then by Lemma 5.2.5, T can be partitioned into P_6 and $\text{wpn}(T) - 2$ edges. Let $\mathcal{N}$ to be the maximal collection of disjoint edges so w' is adjacent to exactly one end of each such edge. Let $P = (P_3, H_2, ..., H_{\text{wpn}(T)})$ where $H_i = P_3$ and for each $j \in [\text{wpn}(T)] \setminus \{1, i\}$, H_j is an edge. As shown in Observation 5.2.56 part (e), there is a partition of P_6 such that either there is a choice of a vertex w which is a part of P_3 and is adjacent to exactly one end of the remaining P_3, or there is a choice of a vertex w which is a part of P_3 and not adjacent to the remaining P_3. Therefore the collection of sets $(\mathcal{P}, \mathcal{H}_2, .., \mathcal{H}_{i-1}, \mathcal{N}, \mathcal{H}_{i+1}, ..., \mathcal{H}_{\text{wpn}(T)})$ is a (P, w', i)-form in Q with property (2) of (b). □

Now we complete the proof of the exact structure for all trees $T \in \mathcal{T}^{\text{npl}}$.

Proof of Theorems 1.1.13, 1.1.15 and 1.1.17. From Observation 5.2.15, to show that the number of bad graphs is much smaller than the number of good graphs it is enough to show that the number of Π-bad graphs is much smaller than the number of Π-good graphs. From Lemmas 5.2.64, 5.2.65, 5.2.67 and 5.2.68 we have the following. Almost all Π-bad graphs G have a set $W(G) \subset V(G)$ such that $|W(G)| \leq 2\,\mathrm{wpn}(T) \cdot (\log n)^3$. Moreover, there is index and $i \in [\mathrm{wpn}(T)]$ such that if $T \in \mathcal{T}^{\mathrm{npl}}$, then $G[\pi_i \setminus W(G)] \in \mathcal{G}_6$ and if $T \in \mathcal{T}^{\mathrm{npl}} \setminus \{P_6\}$, $G[\pi_i \setminus W(G)] \in \mathcal{G}_4$ and if $T \in \mathcal{T}^{\mathrm{npl}} \setminus (\mathcal{S} \cup \{P_6\})$, then $G[\pi_i \setminus W(G)] \in \mathcal{G}_5$, and $\overline{G}[\pi_i \setminus W(G)]$ contains at most at most $\frac{n}{\log\log\log n}$ vertices in components of size at least $(\log n)^2$. Additionally, for all $j \in [\mathrm{wpn}(T)] \setminus \Pi\}$, $G[\pi_j \setminus W(G)]$ is a clique. Let $\mathscr{D}$ be this collection of graphs. Let $\mathcal{F}' = \mathcal{F}(\mathscr{D})$, then by Lemma 5.2.69, $\mathcal{F}'$ is (τ, λ, k)-ordinary $P(T)$-free subsequence for $\tau = \frac{1}{w^2}$, any $\lambda > 0$ small enough and any $k \in \mathbb{N}$. Finally, by Lemma 3.3.14, the number of the remaining Π-bad graphs is much smaller than the number of Π-good graphs. $\square$

The following lemma is needed in the proof in Chapter 6.

Lemma 5.2.70. *Let $T \in \mathcal{T}^{\mathrm{npl}}$ and $n \in \mathbb{N}$ large enough, then almost all T-free graphs G with $|V(G)| = n$ have a partition into $\mathrm{wpn}(T)$ parts $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$ such that without loss of generality $G_1 \in \mathcal{G}_6$ and G_i for $i \geq 2$ is a cliques. Moreover $\overline{G}_1$ contains at least $\frac{n}{(\log n)^2}$ connected components of size at least 2.*

Proof. In this section we showed that almost all T-free graphs are Π-good for some suitable partition Π. We bound the number of good graphs which are Π-good with respect to some partition and have that $\overline{G}[\pi_1]$ contains at most $\frac{n}{(\log n)^2}$ connected components of size at least 2. There are at most $\mathrm{wpn}(T)^n$ ways to obtain a partition Π of $[n]$. Using Lemma 5.2.50, the number of possible graphs on π_1 is at most $2^{4n} \cdot \left(\frac{n}{(\log n)^2}\right)^{\frac{n}{(\log n)^2}}$, but the number of Π-good graphs is at least $\mathrm{Bell}(n)$ which, by Lemma 1.3.2 is at least $\left(\frac{n}{e \ln n}\right)^n$. Hence almost all T-free graphs have the required property. $\square$

Chapter 6

Colouring of Typical T-Free Graphs

In this chapter we prove the asymptotic version of the Gyárfás-Sumner Conjecture. Recall the Gyárfás-Sumner Conjecture that was presented in the introduction.

Conjecture (Gyárfás-Sumner conjecture [30, 54]). *Let H be a forest, then there is a function f_H such that in any H-free graph G, $\chi(G) \leq f_H(\omega(G))$.*

As mentioned, the family $\mathrm{Forb}(H)$ for a forest H, is χ-bounded if and only if all the families $\mathrm{Forb}(T)$ are χ-bounded for each connected component T of H. Hence the above conjecture can be reduced to trees. Also as mentioned, Gyárfás-Sumner conjecture has been proved for the following graphs: paths and stars [30], trees with radius two [35], trees which are subdivided stars [50], trees obtained from trees with radius two by making exactly one subdivision in every edge adjacent to the root [36], "two-legged caterpillars", "double-ended brooms" and a few others [19]. The conjecture is still open in its general form.

We use the structural results we obtained in the previous chapter to show an asymptotic version of the Gyárfás-Sumner conjecture. We recall and prove Theorem 1.0.9.

Theorem (1.0.9). *For every tree T, almost all T-free graphs G have $\chi(G') \leq \mathrm{wpn}(T) \cdot \omega(G')$ for every induced subgraph G' of G. Moreover, for every tree T, almost all T-free graphs G have $\chi(G) = \omega(G)$.*

For the proof we need a definition and the famous Hall's Theorem [13]. Let G be a bipartite graph with a bipartition $V(G) = A \cup B$, that is $G[A]$ and $G[B]$ are stable sets. We say that a matching M in G **saturates** B, if $B \subset \cup_{e \in M} e$.

Theorem 6.0.1 (Hall's theorem [13]). *Let G be a graph with a bipartition $V(G) = A \cup B$. There is a matching which saturates B if and only if for every $W \subseteq B$, $|N(W)| \geq |W|$.*

Recall that $\mathcal{G}_1$ is the family of graphs which are complete multi-partite graphs with parts of size at most 2, and $\mathcal{G}_6$ is the family of graphs which are join of graphs which are disjoint union of a clique and a stable set. Now we prove the main theorem of this chapter.

Proof for Theorem 1.0.9. By Theorems 1.1.3, 1.1.6, and also theorems 1.1.9, 1.1.11, 1.1.13, 1.1.15, and 1.1.17 we can partition the possible $P(T)$-sequences into two classes.

(i) A $\mathcal{P}(T)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ where for each $i \in [\text{wpn}(T)]$, $\mathcal{F}_i \subseteq \mathcal{G}_1$.

(ii) A $\mathcal{P}(T)$-free sequence $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ where the families can be reindexed such that $\mathcal{F}_1 \subseteq \mathcal{G}_6$ and the rest of the families are the sets of all cliques.

Note that in both cases (i) and (ii) each $\mathcal{F}_i \subseteq \text{Forb}(P_4)$, $i \in [\text{wpn}(T)]$, so by Corollary 2.2.5, for each $i \in [\text{wpn}(T)]$, for all $F \in \mathcal{F}_i$ and induced subgraph F' of F, $\chi(F') = \omega(F')$. Also note that in both cases (i) and (ii) for every choice of graphs $(G_1, G_2, ..., G_{\text{wpn}(T)})$ where $G_i \in \mathcal{F}_i$, $i \in [\text{wpn}(T)]$, there is an $i \in [\text{wpn}(T)]$ such that $\omega(G_i) \geq \frac{|V(G_i)|}{2}$.

We first show the first part of the theorem. We show that for every tree T, almost all T-free graphs G have $\chi(G') \leq \text{wpn}(T) \cdot \omega(G')$ for every G' which is an induced subgraph of G. Let G be a typical T-free graph and let $(G_1, G_2, ..., G_{\text{wpn}(T)})$ be its partition such that for each $i \in [\text{wpn}(T)]$, $G_i \in \mathcal{F}_i$ for a $\mathcal{P}(T)$-free sequence as in (i) or (ii). Let G' be an induced subgraph of G and let $(G'_1, G'_2, ..., G'_{\text{wpn}(T)})$ be the corresponding partition of G' such that for each $i \in [\text{wpn}(T)]$, G'_i is an induced subgraphs of G_i. We colour the graphs G'_i iteratively starting from G'_1. We colour each G'_i, $i \in [\text{wpn}(T)]$, with a set of $\omega(G'_i) = \chi(G'_i)$ colours which are different from the colours we already used on graphs $G'_1, G'_2, ..., G'_{i-1}$. Let $w = \max_{i \in [\text{wpn}(T)]} \omega(G_i)$, so in this iterative colouring we used $\sum_{i=1}^{\text{wpn}(T)} \chi(G'_i) = \sum_{i=1}^{\text{wpn}(T)} \omega(G'_i) \leq \text{wpn}(T) \cdot w \leq \text{wpn}(T) \cdot \omega(G')$ colours, as required.

Now we show the second part of the theorem. Let T be a tree, we show that almost all T-free graphs G have $\chi(G) = \omega(G)$. We treat trees with corresponding $P(T)$-free sequences as in class (i) and in class (ii) separately.

Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\mathrm{wpn}(T)})$ be a sequence of class (i) and let G be a T-free graph with a partition $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$ where $G_i \in \mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$. Assume without loss of generality that $\omega(G_1) \geq \omega(G_2) \geq ... \geq \omega(G_{\mathrm{wpn}(T)})$. As before we colour the graphs G_i iteratively starting from G_1. We assign $c_1 := \omega(G_1)$ distinct colours to the vertices $V(G_1)$. It is possible because G_1 is perfect. Let $i \geq 2$, at iteration i, we define an auxiliary bipartite graph A_i, where $V(A_i) = (\mathcal{C}_{i-1}, \mathcal{K}_i)$. The set $\mathcal{C}_{i-1}$ corresponds to the collection of the colour classes in the colouring of $G_1, G_2, ..., G_{i-1}$. We denote the colour of a colour class $C \in \mathcal{C}_{i-1}$ by $c(C)$. Each vertex $K \in \mathcal{K}_i$ corresponds to a connected component in $\overline{G}_i$, note that because we are in case (i), each such component is of size at most 2. The edges $E(A_i)$ are defined as following, an edge $\{C, K = \{v_1, v_2\}\} \in E(A_i)$ if and only if $(N_G(v_1) \cup N_G(v_2)) \cap C = \emptyset$, in particular it means that it is possible to colour the vertices in K with colour $c(C)$ and keep the colouring proper.

We argue that in almost all T-free graphs G, in each iteration $2 \leq i \leq \mathrm{wpn}(T)$, the graph A_i contains a matching which saturates $\mathcal{K}_i$. Let $(G_1, G_2, ..., G_{\mathrm{wpn}(T)})$ be a collection of graphs where $G_i \in \mathcal{F}_i$, $i \in [\mathrm{wpn}(T)]$ for a sequence as in (i). We choose the edges between the parts with probability $\frac{1}{2}$. We want to bound the probability such that there is an $i \in [\mathrm{wpn}(T)]$ where A_i does not contain a matching which saturates $\mathcal{K}_i$. By Hall's Theorem 6.0.1, if such an index $i \in [\mathrm{wpn}(T)]$ exists, then it must be the case that there is a subset $W \subset \mathcal{K}_i$ such that $|N_{A_i}(W)| < |W|$. See a sketch in Figure 6.1.

By our assumption, for each $i \in [\mathrm{wpn}(T)]$, $G_i \in \mathcal{G}_1$, therefore at most 2 vertices in G_i can be in the same colour class. Let $C \in \mathcal{C}_{i-1}$ and $K \in \mathcal{K}_i$, then $|C| \leq 2(i-1)$ and $|K| \leq 2$. Hence the probability for an edge $\{C, K\}$ in A_i is $\left(\frac{1}{2}\right)^{|C|\cdot|K|} \geq \left(\frac{1}{2}\right)^{2(i-1)\cdot 2} = \frac{1}{2^{4(i-1)}}$. By our assumptions $|\mathcal{C}_{i-1}| \geq |\mathcal{K}_i|$, therefore the expected number of neighbours in A_i of a vertex $v \in V(A_i)$ is at least $\frac{|\mathcal{K}_i|}{2^{4(i-1)}} \geq \frac{n}{2\,\mathrm{wpn}(T)\cdot 2^{4(i-1)}}$. Using Chernoff bound 1.3.6, the probability that there is an index $i \in [\mathrm{wpn}(T)]$ and a vertex $v \in V(A_i)$ such that $|N_{A_i}(v)| \leq \frac{n}{2^2\,\mathrm{wpn}(T)\cdot 2^{4(i-1)}}$, is at most

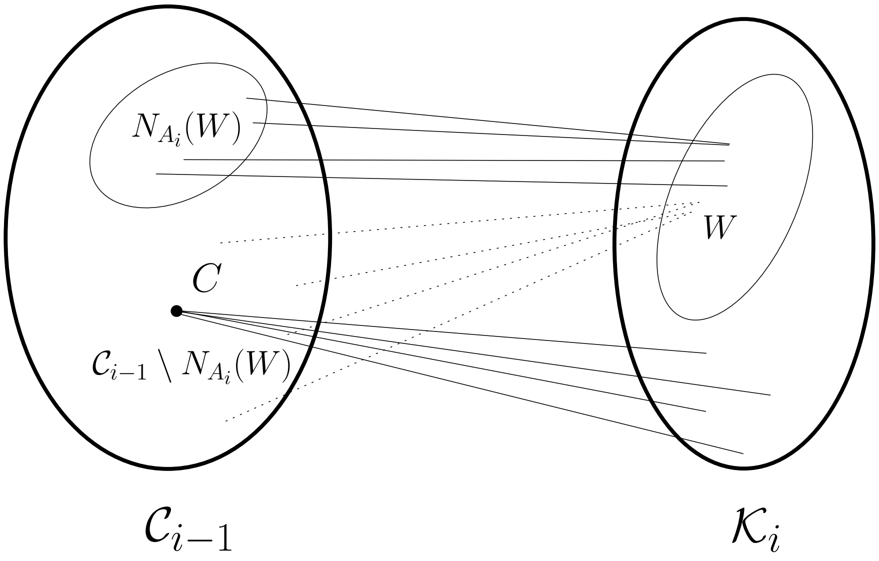


Figure 6.1: Sketch of the graph A_i.

$\operatorname{wpn}(T) \cdot n \cdot e^{-\frac{n}{2^3 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}}$. Therefore the number of such graphs is much smaller than the number of T-free graphs without such a vertex. Hence we can make the following assumptions about the graph we consider. The size of the neighbourhood of W in A_i is at least the size of the neighbourhood of some vertex $w \in W$, that is $|N_{A_i}(W)| \geq |N_{A_i}(w)| \geq \frac{n}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}$. Moreover by our assumptions $|N_{A_i}(W)| < |W|$ and $|\mathcal{C}_{i-1}| \geq |\mathcal{K}_i|$, so $\mathcal{C}_{i-1} \setminus N_{A_i}(W) \neq \emptyset$. Let $C \in \mathcal{C}_{i-1} \setminus N_{A_i}(W)$, then $|N_{A_i}(C)| \geq \frac{n}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}$. By the choice of C, $N_{A_i}(C) \cap W = \emptyset$, so $|\mathcal{K}_i \setminus W| \geq |N_{A_i}(C)| \geq \frac{n}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}$. We summarize,

$$\frac{n}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}} \leq |N_{A_i}(W)| < |W| \leq |\mathcal{K}_i| - |\mathcal{K}_i \setminus W| \leq \left(1 - \frac{1}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}\right) n.$$

Moreover,

$$\frac{n}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}} \leq |\mathcal{C}_{i-1}| - |W| \leq |\mathcal{C}_{i-1} \setminus N_{A_i}(W)| \leq \left(1 - \frac{1}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}\right) n.$$

Therefore we can conclude that the probability for the event that there is an $i \in [\operatorname{wpn}(T)]$ such that in A_i there is a set $W \subset \mathcal{K}_i$ where $|N_{A_i}(W)| < |W|$ is at most the probability that there are no edges between W and $\mathcal{C}_{i-1} \setminus N_{A_i}(W)$ which is,

$$2^n \cdot \left(\frac{1}{2}\right)^{\left(\frac{n}{2^2 \operatorname{wpn}(T) \cdot 2^{4(i-1)}}\right)^2}$$

which means that in almost all T-free graphs G, such a set W does not exist. Therefore we can colour $V(G)$ with at most $\omega(G_1) \leq \omega(G)$ colours.

Now we consider the remaining case for the $\mathcal{P}(T)$-free sequence. Let $(\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_{\text{wpn}(T)})$ be a sequence of class (ii) and let G be a T-free graph with a partition $(G_1, G_2, ..., G_{\text{wpn}(T)})$ where $G_i \in \mathcal{F}_i$, $i \in [\text{wpn}(T)]$. As mentioned, we have that $G_1 \in \mathcal{G}_6$, we describe how to partition $V(G_1)$ into stable sets of at most 2 vertices. Let $K_1, K_2, ..., K_k$ for some $k \in \mathbb{N}$ be the connected components in $\overline{G}_1$, and let S_i, C_i be the partition of K_i, $i \in [k]$, into the stable set and the clique. If $|V(C_i)| \geq |V(S_i)|$, then we match every vertex from S_i to a vertex in C_i, and every such matched pair of vertices is a part in the partition of $V(G_1)$. If $|V(C_i)| < |V(S_i)|$, then we match $|V(C_i)|$ from S_i to $|V(C_i)|$ vertices of C_i and again every such matched pair of vertices is a part in the partition of $V(G_1)$, the rest $|V(S_i)| - |V(C_i)|$ vertices from S_i we partition arbitrary into $\lfloor \frac{|V(S_i)| - |V(C_i)|}{2} \rfloor$ pairs and maybe one more singleton part. By Lemma 5.2.70, $\overline{G}_1$ contains at least $\frac{n}{(\log n)^2}$ connected components, therefore $\omega(G_1) \leq |V(G_1)| - \frac{n}{(\log n)^2}$ and therefore there is $i \geq 2$ such that $\omega(G_i) \geq \omega(G_1)$. We proceed as in the previous case and this gives us the required colouring. $\square$

Chapter 7

Summary and Future Work

In this thesis we have studied the structure of almost all T-free graphs for any tree T. In Chapter 2 we reviewed some of the results regarding the structure of H-free graphs both for any graph H and for some specific graphs H. In Chapter 3, we proved weaker results regarding the structure of almost all H-free graphs. Moreover, we developed some general tools which can be applied to nearly all graphs H. In Chapter 4, we gave a formula for the value of the witnessing partition number of any bipartite graph H. In Chapter 5, we reproved the result of Balogh and Butterfield [5] regarding the structure of almost all H-free graphs where H is a critical graph. Moreover, we proved the exact structure of almost all T-free graphs for any tree T. Finally, in Chapter 6, we used our result regarding the structure of almost all T-free graphs to show that almost all T-free graphs are χ-bounded, which is an asymptotic version of the Gyárfás-Sumner Conjecture [30, 54]. In the following we list some possible future directions for the thesis topic.

7.1 Reed-Scott Conjecture

We recall the statement of Reed-Scott conjecture.

Conjecture (Reed-Scott [47], 1.1.1)**.** *For every graph H, almost all H-free graphs G have a $P(H)$-free partition.*

We proved the above conjecture for trees. The starting point of our proof was understanding the set $P(T)$ of partitions of T. For the rest of the proof we did not use any additional properties of the trees besides the structure of the partitions in $P(T)$. Note that we divided the proof into the different subfamilies of trees because the trees in the different subfamilies have somewhat different partitions into smaller graphs. Due to the fact that the only property we used is the structure of the partitions in $P(T)$, our proof can be applied to some graphs H which have the same minimal partitions in $P(H)$ as in $P(T)$ for some tree T.

A very natural next step is to try to prove Reed-Scott conjecture for bipartite graphs B which have a set of partitions $P(B)$ different from the ones which were already considered. As proved in Theorem 4.0.1, the value of the witnessing partition number of a (non-complete) bipartite graph can be computed in the same way as the witnessing partition number of a tree. Therefore the first step of the proof would be to find all the partitions in $P(B)$. Knowing those partitions allows us to obtain the $P(B)$-free sequences and understand which graphs are extendable (in any of the definitions of extendability). Once we have all the above, one way to proceed with the proof is to define subfamilies of B-free graphs which do not have a $P(H)$-free partition and to show that each of those subfamilies is much smaller than the number of graphs which have a $P(H)$-free partition. It would be useful to define those families in a way that allows us to gain a better understanding of the structure of almost all B-free graphs. This in a very general terms was what we did for the case of trees.

After the case of bipartite graphs, it would be interesting to consider the case of all triangle-free graphs. As before, one way to prove Conjecture 1.1.1 for such graphs H, is by first finding the value of the witnessing partition number of H, the set of possible partitions $P(H)$ and then proceed obtaining better and better structure until the point where one shows that almost all H-free graphs have a $P(H)$-free partition.

In the opposite direction, it could be interesting to try to disprove Conjecture [47]. To this end, one possible approach is to try to find a graph H such that almost all H-free graphs have an almost $P(H)$-free partition. In other words, a graph H where there are many H-free graphs G and there is a set of vertices $B \subseteq V(G)$ such that $G[V(G) \setminus B]$ has a $P(H)$-free partition, but there is a way to add the vertices B so there is still no induced copy of H in G.

7.2 Applying the Techniques Further

The techniques which were developed in this thesis could be applied to a range of families of graphs which can be characterized by forbidding some collection of induced subgraphs.

For example, recently with János Pach, we have applied the ideas from the thesis and [47] to string graphs [42]. A **string graph** is the intersection graph of a family of continuous arcs in the plane. We showed that the vertex set of almost all string graphs on n vertices can be partitioned into four sets, such that three of them are cliques and the last set is a disjoint union of two cliques ($n \to \infty$). The intersection graph of a family of plane convex sets is a string graph, but not all string graphs can be obtained in this way. Together with some additional arguments we could show that almost all string graphs on n vertices are intersection graphs of plane convex sets. This result also verified a conjecture by Janson and Uzzell [31].

The main lemma that allowed us to apply our techniques to string graphs was that the graph which is a subdivision of K_5, together with some additional edges connecting the vertices of degree 2 in the subdivision, is not a string graph.

As a future work it would be of interest to try to apply our techniques to additional families of graphs.

7.3 Graphs without $U(k)$

From Theorems 3.2.1 and 3.2.2 together with 3.3.2, we know that for any graph H, almost all H-free graphs G contain a set $Z(G) \subseteq V(G)$ such that $G[V(G) \setminus Z(G)]$ has a partition into $\mathrm{wpn}(H)$ parts where each of the parts does not contain a copy of $U(k)$. Moreover $|Z(G)| \leq |V(G)|^{1-\varepsilon}$ for some $\varepsilon > 0$. Therefore an understanding of graphs without a copy of

$U(k)$ can lead to a better understanding of the structure of almost all H-free graphs for all graphs H.

Here is a simple result for $k = 2$. A **threshold graph** is a graph which does not contain an induced $2K_2, P_4$ or C_4. A **bipartite threshold graph** is a bipartite graph which does not contain an induced $2K_2$. A **half-graph** is a bipartite graph with the bipartition $(\{a_1, a_2, ..., a_k\}, \{b_1, b_2, ..., b_k\})$ in which a_j is adjacent to b_j if and only if $i + j \geq k + 1$. An *alternating blow up of a* P_5 is a graph which is a $P_5 = v_1, v_2, v_3, v_4, v_5$ where the vertices v_1, v_3, v_5 are blown up into stable sets. The **bull** graph is a graph which is combined of $P_4 = v_1, v_2, v_3, v_4$ together with a vertex u which is adjacent to v_2 and v_3.

Theorem 7.3.1 (Norin and Yuditsky)**.** *Let G be a graph without a copy of $U(2)$, then the following is true.*

- *If G is not C_5-free then $G = C_5$.*
- *If G is not P_5-free, then G is an alternating blow up of a P_5.*
- *If G is $\{P_5, \overline{P_5}, C_5\}$-free. Then either G or $\overline{G}$ is a disjoint union of graphs $G_1, G_2, ..., G_k$ for some $k \in \mathbb{N}$ and each G_i, $i \in [k]$, is one of the following graphs,*
  1. *threshold graph,*
  2. *bipartite threshold graph,*
  3. *graph obtained from the bull graph by substituting each vertex of the P_4 with either a clique or a stable set, and substituting u with a threshold graph,*
  4. *if $k = 1$, graph obtained from P_4 by substituting each vertex with either a clique or a stable set, and if $k \geq 2$, graph obtained from P_4 by substituting vertices v_2 and v_3 with stable sets,*
  5. *if $k = 1$, graph obtained from the half-graph by substituting each vertex with a threshold graph, and if $k \geq 2$, graph obtained from the half-graph by substituting each vertex with a threshold graph, where additionally for every edge $\{v_1, v_2\}$ of*

the half-graph, either both v_1 *and* v_2 *are stable sets, or one of them is threshold graph and the other is a clique.*

It would be of interest to generalize the above result to $U(k)$ for $k = 3$ or greater.

Bibliography

[1] V.E. Alekseev. Range of values of entropy of hereditary classes of graphs. Diskret. Mat. 148-157 (1992).

[2] N. Alon, J. Balogh, B. Bollobás and R. Morris. The structure of almost all graphs in a hereditary property. J. Combin. Theory B 101: 85-110 (2011).

[3] K. Appel and W. Haken. Every planar map is four colorable. Part I: Discharging. Illinois Journal of Mathematics, 21(3): 429-490 (1977).

[4] M. Axenovich, R. Martin, A. Kezdy. Editing distance of graphs. Journal of Graph Theory, Volume 58, Number 2: 123 - 138 (2008).

[5] J. Balogh and J. Butterfield. Excluding induced subgraphs: Critical graphs. Random Structures and Algorithms, 38(1-2): 100-120 (2011).

[6] J. Balogh, B. Bollobás and M. Simonovits. The number of graphs without forbidden subgraphs. J. Comb. Theory Ser. B, 91(1): 1–24 (2004).

[7] J. Balogh, B. Bollobás, and D. Weinreich. The speed of hereditary properties of graphs. J Combin Theory Ser B 79: 131-156 (2000).

[8] J. Balogh, R. Morris, W. Samotij, Independent sets in hypergraphs. J. Amer. Math. Soc 28: 669-709 (2015).

[9] J. Balogh, R. Morris, W. Samotij and L. Warnke. The typical structure of sparse K_{r+1}-free graphs. Trans. Amer. Math Soc., 368: 6439-6485 (2016).

[10] D. Berend and T. Tassa. Improved bounds on Bell numbers and on moments of sums of random variables. Probability and Mathematical Statistics 30(2): 185-205 (2010).

[11] C. Berge. Perfect graphs. Six Papers on Graph Theory. Calcutta: Indian Statistical Institute: 1-21 (1963).

[12] B. Bollobás and A. Thomason, Hereditary and monotone properties of graphs. The mathematics of Paul Erdős II 14: 70-78 (1997).

[13] J.-A. Bondy and U.S.R Murty. Graph theory. Graduate texts in mathematics, Springer, (2007).

[14] N. G. de Bruijn. Asymptotic Methods in Analysis. Dover, New York, NY: 103-108 (1958).

[15] A. Cayley. A theorem on trees. Quart. J. Pure Appl. Math. 23: 376-378 (1889); Collected Mathematical Papers Vol. 13, Cambridge University Press: 26-28, (1897).

[16] H. Chernoff. A note on an inequality involving the normal distribution. Ann. Probab., 9: 533-535 (1981).

[17] M. Chudnovsky. The Erdős-Hajnal Conjecture—A Survey. `https://arxiv.org/abs/1606.08827`

[18] M. Chudnovsky, N. Robertson, P. Seymour and R. Thomas. The strong perfect graph theorem. Annals of Mathematics, 164(1): 51-229 (2006).

[19] M. Chudnovsky, A. Scott and P. Seymour. Induced subgraphs of graphs with large chromatic number. XII. Two-legged caterpillars. `https://web.math.princeton.edu/~pds/papers/cuttingedge/paper.pdf`.

[20] D.G. Corneil and H. Lerchs and L. Stewart Burlingham. Complement reducible graphs. Discrete Applied Mathematics (3): 163 - 174 (1981).

[21] P. Erdős. Graph theory and probability. Canad. J. Math. 11: 34-38 (1959).

[22] P. Erdős. On some new inequalities concerning extremal properties of graphs. P. Erdos and G. Katona, eds., Theory of Graphs: 77-81 (1968).

[23] P. Erdős. Some recent results on extremal problems in graph theory. P. Rosenstiehl, ed., Theory of Graphs: 117-123 (1976).

[24] P. Erdős, P. Frankl and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. Graphs Combin. 2: 113-121 (1986).

[25] P. Erdős and A. Hajnal. Ramsey-type theorems. Discrete Appl. Math., 25(1-2): 37-52 (1989).

[26] P. Erdős, D. J. Kleitman and B. L. Rothschild. Asymptotic enumeration of K_n-free graphs. International Colloquium on Combinatorial Theory, Atti dei Convegni Lincei 17(2): 19-27 (1976).

[27] P. Erdős and M. Simonovits. A limit theorem in graph theory. Studia Sci. Math. Hungar. 1: 51-57 (1966).

[28] P. Erdős and A.H. Stone. On the structure of linear graphs. Bull. Amer. Math. Sot. 52: 1089-1091 (1946).

[29] P. Erdős and G. Szekeres. A combinatorial problem in geometry. Compositio Math. 2: 463-470 (1935).

[30] A. Gyárfás. Problems from the world surrounding perfect graphs. Zastosow Ania Matematyki, Applicationes Mathematicae, XIX, 3-4: 413-441 (1987).

[31] S. Janson and A. Uzzell. On String Graph Limits and the Structure of a Typical String Graph. J. Graph Theory 84: 386-407 (2017).

[32] D.E. Knuth. The Art of Computer Programming, Volume 3, 2nd ed. Addison-Wesley: 73-75 (1997).

[33] R. Kang, C. McDiarmid, B. Reed and A. Scott. For most graphs H, most H-free graphs have a linear homogeneous set. Random Structures and Algorithms 45(3): 343-361 (2014).

[34] J. Kim, D. Kühn, D. Osthus, T. Townsend. Forbidding induced even cycles in a graph: typical structure and counting. `http://arxiv.org/abs/1507.04944`.

[35] H. A. Kierstead and S. G. Penrice. Radius two trees specify χ-bounded classes. Journal of Graph Theory 18(2): 119-129 (1994).

[36] H. A. Kierstead and Y. Zhu. Radius Three Trees in Graphs with Large Chromatic Number. SIAM J. Discrete Math. 17(4): 571-581 (2004).

[37] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi. The regularity lemma and its applications in graph theory. Theoretical Aspects of Computer Science: 84-112 (2000).

[38] Ph.G. Kolaitis, H.J. Prömel and B.L. Rothschild. $K_{\ell+1}$-free graphs: asymptotic structure and a 0 - 1 law. Trans. Amer. Math. Soc. 303: 637-671 (1987).

[39] M. Loebl, B. Reed, A. Scott, A. Thomason and S. Thomassé. Almost all H-free graphs have the Erdős-Hajnal property. An Irregular Mind (Szemerédi is 70): 405-414 (2010).

[40] J. Mycielski. Sur le coloriage des graphes. Colloq. Math. 3: 161-162 (1955).

[41] D. Osthus, H. J. Prömel and A. Taraz. For Which Densities are Random Triangle-Free Graphs Almost Surely Bipartite? Combinatorica 23(1): 105–150, (2003).

[42] J. Pach, B. Reed and Y. Yuditsky. Almost all string graphs are intersection graphs of plane convex sets. To be submitted.

[43] H. J. Prömel and A. Steger. Excluding induced subgraphs: Quadrilaterals. Random Structures and Algorithms 2: 55-71 (1991).

[44] H. J. Prömel and A. Steger. Excluding induced subgraphs III: A general asymptotic. Random Structures and Algorithms 3(1): 19-31 (1992).

[45] H. J. Prömel and A. Steger. Almost all Berge graphs are perfect. Combinatorics, Probability and Computing 1: 53-79 (1992).

[46] H. J. Prömel and A. Steger. Excluding induced subgraphs II: Extremal graphs. Discrete Applied Mathematics 44(1-3): 283-294 (1993).

[47] B. Reed, A. Scott. The Global Structure of a Typical Graph without H as an induced subgraph when H is a cycle. To be submitted.

[48] D. Saxton and A. Thomason. Hypergraph containers. Inventiones mathematicae, 201(3): 925-992 (2015).

[49] E. Scheinerman and J. Zito. On the size of hereditary classes of graphs. J Combin Theory Ser B 61: 16-39 (1994).

[50] A. D. Scott. Induced trees in graphs of large chromatic number. J. Graph Theory 24: 297-311 (1997).

[51] D. Seinsche On a property of the class of n-colorable graphs. J. Combin. Theory Ser. B 16: 191-193 (1974).

[52] M. Simonovits. A method for solving extremal problems in graph theory. P. Erdos and G. Katona, eds., Theory of Graphs: 279-310 (1968).

[53] M. Simonovits. Extremal graph problems with symmetrical extremal graphs. Additional chromatic conditions. Discrete Math. 7: 349-376 (1974).

[54] D.P. Sumner. Subtrees of a graph and chromatic number. The Theory and Applications of Graphs, John Wiley & Sons, New York: 557-576 (1981).

[55] E. Szemerédi. Regular partitions of graphs. Colloques Internationaux C.N.R.S. No 260 - Problemes Combinatoires et Theorie des Graphes, Orsay: 399-401 (1976).

[56] P. Turán. On an extremal problem in graph theory (in Hungarian). Mat. Fiz. Lupok, 48: 436-452 (1941).

Index