1-1-2018

Problems in Extremal Graph Theory

Kamil Popielarz

PROBLEMS IN EXTREMAL GRAPH THEORY

by

Kamil Krzesimir Popielarz

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Mathematical Sciences

The University of Memphis

May 2018

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Béla Bollobás, for his guidance and support, and for the opportunity of being his student. Being around him and all his students in the wonderful places like Cambridge, Lucca, or Rio de Janeiro, made me a better mathematician in every possible aspect.

I would also like to thank my committee members: Professor Paul Balister, Professor James T. Campbell, Professor Ebenezer Olúṣẹ́gun George, and Professor David Grynkiewicz, for devoting their precious time to me.

I am very grateful to Mrs. Tricia Simmons for her hard work and patience, and for always finding time to answer my silly questions. And to Mrs. Gabriella Bollobás for always inviting us to her house, for all the lively conversations, and memorable parties.

Big thanks to my collaborators (in reverse alphabetical order): Richard Snyder, Julian Sahasrabudhe, Gábor Mészáros, David Lewis, Teeradej Kittipassorn, and António Girão, for many hours of interesting and joyful discussions.

My life as a graduate student would have been much harder if it hadn't been for many friends I made in Memphis and Cambridge. To avoid the risk of omission I am not going to name them here. Thanks to you all.

And last, but not least, I would like to thank my family, especially my parents, for their constant support. None of this would be possible without them.

Dziękuję

ABSTRACT

Popielarz, Kamil Ph.D. The University of Memphis, May 2018. Problems in Extremal Graph Theory. Major Professor: Béla Bollobás, Ph.D.

This dissertation consists of 6 chapters concerning a variety of topics in extremal graph theory.

Chapter 1 is dedicated to the results in the papers with António Girão, Gábor Mészáros, and Richard Snyder [47, 44]. We say that a graph is path-pairable if for any pairing of its vertices there exist edge disjoint paths joining the vertices in each pair. We study the extremal behavior of maximum degree and diameter in some classes of path-pairable graphs. In particular we show that a path-pairable planar graph must have a vertex of linear degree.

In Chapter 2 we present a joint work with António Girão and Teeradej Kittipassorn [46]. Given graphs G and H, we say that a graph F is H-saturated in G if F is H-free subgraph of G, but addition of any edge from $E(G)$ to F creates a copy of H. Here we deal with the case when G is a complete k-partite graph with n vertices in each class, and H is a complete graph on r vertices. We prove bounds, which are tight for infinitely many values of k and r, on the minimal number of edges in a H-saturated graph in G, for this choice of G and H, answering questions of Ferrara, Jacobson, Pfender, and Wenger; and generalizing a result of Roberts.

Chapter 3 is about a joint paper with António Girão and Teeradej Kittipassorn [43]. A coloring of the vertices of a digraph D is called *majority coloring* if no vertex of D receives the same color as more than half of its outneighbours. This was introduced by van der Zypen in 2016. Recently, Kreutzer, Oum, Seymour, van der Zypen, and Wood posed a number of problems related to this notion: here we solve several of them.

In Chapter 4 we present a joint work with António Girão [45]. We show that given any integer k there exist functions $g_1(k), g_2(k)$ such that the following holds. For any graph G with maximum degree Δ one can remove fewer than $g_1(k)\sqrt{\Delta}$ vertices from G so that the remaining graph H has k vertices of the same degree at least $\Delta(H) - g_2(k)$. It is an

approximate version of conjecture of Caro and Yuster; and Caro, Lauri, and Zarb, who conjectured that $g_2(k) = 0$.

Chapter 5 concerns results obtained together with Kazuhiro Nomoto, Julian Sahasrabudhe, and Richard Snyder. We study a graph parameter, the *graph burning number*, which is supposed to measure the speed of the spread of contagion in a graph. We prove tight bounds on the graph burning number of some classes of graphs and make a progress towards a conjecture of Bonato, Janssen, and Roshanbin about the upper bound of graph burning number of connected graphs.

In Chapter 6 we present a joint work with Teeradej Kittipassorn. We study the set of possible numbers of triangles a graph on a given number of vertices can have. Among other results, we determine the asymptotic behavior of the smallest positive integer m such that there is no graph on n vertices with exactly m copies of a triangle. We also prove similar results when we replace triangle by any fixed connected graph.

TABLE OF CONTENTS

CHAPTER 1

ON SOME PARAMETERS OF PATH-PAIRABLE GRAPHS

A graph is *path-pairable* if for any pairing of its vertices there exist edge-disjoint paths joining the vertices in each pair. In this chapter we study the extremal behavior of two graph parameters, maximum degree and diameter, in some classes of path-pairable graphs. We show that any n-vertex path-pairable planar graph must contain a vertex of degree linear in n. We also obtain sharp bounds on the maximum possible diameter of path-pairable graphs which either have a given number of edges, or are c-degenerate. This work is joint with António Girão, Gábor Mészáros, and Richard Snyder.

1.1 Introduction

Path-pairability is a graph theoretical notion that emerged from a practical networking problem. This notion was introduced by Csaba, Faudree, Gyárfás, Lehel, and Schelp [26], and further studied by Faudree, Gyárfás, and Lehel [34, 40, 36], and by Kubicka, Kubicki and Lehel [60]. Given a fixed integer k and a simple undirected graph G on at least $2k$ vertices, we say that G is *k-path-pairable* if, for any pair of disjoint sets of distinct vertices $\{x_1,\dots,x_k\}$ and $\{y_1,\dots,y_k\}$ of G, there exist k edge-disjoint paths $P_1,P_2,\dots,P_k$, such that P_i is a path from x_i to y_i, $1 \le i \le k$. The problems of finding k edge(vertex)-disjoint paths routing some prescribed pairs of vertices in a graph is a well-known problem in algorithmic graph theory and combinatorial optimization (see the surveys[40, 41, 77]). Recently, for a fixed integer k, Kawarabayashi, Kobayashi and Reed [55] constructed a $O(n^2)$ time algorithm which for any graph G on n vertices either finds such k vertex-disjoint paths or concludes that no such paths exist. As a corollary they obtained a $O(n^2)$ time algorithm for the edge-disjoint paths problem. This improved upon the seminal work of Robertson and Seymour [73], which initially gave a $O(n^3)$ time algorithm for the vertex-disjoint path problem. Note that the problem of finding edge(vertex)-disjoint paths between an unbounded number of prescribed pairs of vertices is known to be NP-complete, even when restricted to planar graphs [67].

The concept of k-path-pairability is closely related to the well-studied notions of *k-linkedness* and *k-weak-linkedness*. A graph is said to be k-(weakly)linked if for any choice $\{s_1,\ldots,s_k,t_1,\ldots,t_k\}$ of $2k$ vertices (not necessarily distinct) there are vertex(edge) internally disjoint paths $P_1,\ldots,P_k$with P_i joining s_i to t_i, $1 \leq i \leq k$. While any k-(weakly)linked graph is $(2k-1)$-vertex connected (k-edge connected), the same need not hold for k-path-pairable graphs. Observe that the *stars* S_{2k} ($k \geq 1$) are k-path-pairable and yet have very low edge density and edge connectivity. On the other hand, a result of Bollobás and Thomason [15] shows that if G is a $2k$-connected graph with average degree at least $22k$ then G is k-linked. This was later improved by Thomas and Wollan [76] who showed that a $2k$-connected graph with average degree at least $10k$ is necessarily k-linked. In the context of weakly-linked graphs, a theorem of Hirata, Kubota and Saito [51] states that a $(2k+1)$-edge connected graph is $(k+2)$-weakly-linked for $k \geq 2$. A few years later, Huck [52] showed that any $(k+2)$-edge-connected graph is k-weakly-linked.

A k-path-pairable graph on $2k$ or $2k+1$ vertices is simply said to be *path-pairable*.

It is fairly easy to construct path-pairable graphs on n vertices (n even) with maximum degree linear in n and/or small (constant) diameter. For example, complete graphs K_{2n} and complete bipartite graphs $K_{m,n}$ are path-pairable for all choices of $m,n \in \mathbb{N}$ with $m+n$ even, $m \neq 2, n \neq 2$.

It is slightly more challenging to construct an infinite family of path-pairable graphs where the maximum degree grows sublinearly or the diameter grows with the number of vertices. We shall now describe such a family. Let K_t be the complete graph on t vertices and let K_t^q be constructed from K_t by attaching $q-1$ leaves to each of the original vertices of K_t. This family was introduced by Csaba, Faudree, Gyárfás, Lehel, and Schelp [26], who also proved that K_t^q is path-pairable as long as $t \cdot q$ is even and $q \leq \left\lfloor \frac{t}{3+2\sqrt{2}} \right\rfloor$. The bound on q has been recently improved to $\approx \frac{1}{3}t$ [49]. Observe that $n = |V(K_t^q)| = t \cdot q$ and

$\Delta(K_t^q) = t + q - 2 = O(\sqrt{n})$ when $q = \Omega(t)$. Additional path-pairable constructions with maximum degree $c\sqrt{n}$ can be found in [60] and [65]. A construction of path-pairable graphs with unbounded diameter, which is based on a blow-up of a path, is presented in a later section.

The following result due to Faudree [35] shows that the maximum degree of a path-pairable graph has to grow with the order of the graph.

Theorem 1.1. If G is path-pairable on n vertices with maximum degree Δ, then $n \leq 2\Delta^{\Delta}$.

Letting $\Delta_{\min}(n) := \min\{\Delta(G) : G \text{ is a path-pairable graph on } n \text{ vertices}\}$, this result is equivalent to

$$\Delta_{\min}(n) \geq c_1 \frac{\log n}{\log\log n},$$

for some constant c_1. To date, the best known upper bound on $\Delta_{\min}(n)$ is due to Győri, Mezei, and Mészáros, exhibiting a path-pairable graph with maximum degree $\Delta \approx 5.5 \cdot \log n$ [48]. In summary, we have the following general asymptotic bounds on $\Delta_{\min}(n)$:

$$c_1 \frac{\log n}{\log\log n} \leq \Delta_{\min}(n) \leq c_2 \log n.$$

The maximum diameter of arbitrary path-pairable graphs was investigated by Mészáros [65] who proved that $d(n) \leq 6\sqrt{2}\sqrt{n}$.

Recall that the star $K_{1,n-1}$ is path-pairable. This is simply due to the presence of a vertex of large degree. Are there properties we may impose on a general path-pairable graph to force a vertex of large degree, say, linear in n? Along these lines, Faudree, Gyárfás and Lehel [36] proved that an n-vertex path-pairable graph with maximum degree at most $n-2$ must have at least $3n/2 - \log n - c$ edges, for some absolute constant c. Instead of simply imposing a condition on the number of edges, we wished to determine whether or not a structural property like *planarity* would be enough to force a vertex of linear degree in a path-pairable graph. To formulate this precisely, let us define $\Delta_{\min}^{p}(n)$ to be

$$\min\{\Delta(G) : G \text{ is a path-pairable planar graph on } n \text{ vertices}\}.$$

Our problem, then, is to determine whether or not $\Delta^p_{\min}(n) = \Theta(n)$. We first note that a simple application of the Planar Separator Theorem of Lipton and Tarjan [63] shows that every path-pairable planar graph on n vertices must contain a vertex of degree at least $c\sqrt{n}$. Indeed, if G is such a graph, then the Separator Theorem allows us to partition $V(G)$ into three sets S, A, B, where $|S| = O(\sqrt{n})$, $|A| \leq |B| \leq 2n/3$, and there are no edges between A and B. Now, while path-pairable graphs G need not be highly connected or edge connected, they must satisfy certain connectivity-like conditions. More precisely, they must satisfy the *cut-condition*: for every subset $X \subset V(G)$ of size at most $n/2$, there are at least $|X|$ edges between X and $V(G) \setminus X$. Note that the cut-condition is not sufficient to guarantee path-pairability; see [66] for additional details. Accordingly, since $n/4 < |A| < n/2$ and there are no edges between A and B, the cut-condition implies that there are at least $|A|$ edges between A and S. We therefore obtain a vertex in S of degree $\Omega(\sqrt{n})$.

Our main theorem, which we state below, shows that we can do much better than this. Namely, every path-pairable planar graph must have a vertex of *linear* degree.

Theorem 1.2. There exists $c \geq 10^{-10^{10}}$ such that if G is a path-pairable planar graph on n vertices then $\Delta(G) \geq cn$.

We have not made an attempt to optimize the constant c obtained in the proof. The value we give is surely far from the truth.

In the other direction, there are easy examples of path-pairable planar graphs with very large maximum degree; for example, consider the star $K_{1,n-1}$. Our second result finds an infinite family of path-pairable planar graphs with smaller (but of course still linear) degree.

Theorem 1.3. There exist path-pairable planar graphs G on n vertices with $\Delta(G) = \frac{2}{3}n$.

Combining Theorems 1.2 and 1.3, we have that

$$10^{-10^{10}} n \le \Delta^p_{\min}(n) \le \frac{2}{3}n.$$

For the diameter problem, for a family of graphs $\mathscr{G}$ let us define $d(n,\mathscr{G})$ as follows:

$$d(n,\mathscr{G}) = \max\{d(G) : G \in \mathscr{G} \text{ and } G \text{ is path-pairable on } n \text{ vertices}\}.$$

When $\mathscr{G}$ is the family of path-pairable graphs, we shall simply write $d(n)$ instead of $d(n,\mathscr{G})$.

The maximum diameter of arbitrary path-pairable graphs was investigated by Mészáros [65] who proved that $d(n) \le 6\sqrt{2}\sqrt{n}$. Our next aim in this chapter is to investigate the maximum diameter of path-pairable graphs when we impose restrictions on the number of edges and on how the edges are distributed. To state our results, let us denote by $\mathscr{G}_m$ the family of graphs with at most m edges. The following result determines $d(n,\mathscr{G}_m)$ for a certain range of m.

Theorem 1.4. If $2n \le m \le \frac{1}{4}n^{3/2}$ then

$$\sqrt[3]{\frac{1}{2}m - n} \le d(n,\mathscr{G}_m) \le \sqrt[3]{300m}.$$

We remark that we actually prove a slightly more general bound $d(n,\mathscr{G}_m) \le \max\left\{\frac{6m}{n}, \sqrt[3]{300m}\right\}$ which holds for m in any range but, when $m \ge \sqrt{2}n^{3/2}$, the bound obtained by Mészáros [65] is sharper. Determining the behavior of the maximum diameter among path-pairable graphs on n vertices with fewer than $2n$ edges remains an open problem. In particular, we do not know if the maximum diameter in this range must be bounded (see Section 1.3.4).

Following this line of research, it is very natural to consider the problem of determining the maximum attainable diameter for other classes of graphs. For example, what is the behavior of the maximum diameter of path-pairable *planar* graphs? Although we could not give a satisfactory answer to this particular question, we were able to do so for graphs which are *c-degenerate*. As usual, we say that an n-vertex graph G is *c-degenerate* if there exists an ordering $v_1, \ldots, v_n$ of its vertices such that $|\{v_j : j > i, v_iv_j \in E(G)\}| \leq c$ holds for all $i = 1, 2, \ldots, n$. We let $\mathscr{G}_{c\text{-deg}}$ denote the family of c-degenerate graphs. Clearly all c-degenerate graphs have a linear number of edges, so Theorem 1.4 implies that $d(n, \mathscr{G}_{c\text{-deg}}) = O(\sqrt[3]{n})$. However, as the next result shows, this bound is far from the truth.

Theorem 1.5. Let $c \geq 5$ be an integer. Then

$$(4+o(1))\frac{\log(n)}{\log(\frac{c}{c-2})} \leq d(n, \mathscr{G}_{c\text{-deg}}) \leq (12+o(1))\frac{\log(n)}{\log(\frac{c}{c-2})}$$

as $n \to \infty$.

We remark that we have not made an effort to optimize the constants appearing in the upper and lower bounds of Theorems 1.4 and 1.5.

1.2 Maximum degree of path-pairable planar graphs

1.2.1 The Construction

Our aim in this section is to prove Theorem 1.3, which we restate here for convenience.

Theorem 1.3. There exist path-pairable planar graphs G on n vertices with $\Delta(G) = \frac{2}{3}n$.

Proof. Let G be a graph on $n = 6k$ vertices with vertex set $V(G) = A \cup B \cup C \cup \{x_{AB}, x_{BC}, x_{CA}\}$ where $|A| = |B| = |C| = 2k-1$, and x_{AB}, x_{BC}, x_{CA} denote three additional vertices forming a triangle such that x_{AB}, x_{BC}, x_{CA} are joined to every vertex in $A \cup B$, $B \cup C$, and $C \cup A$, respectively, and A, B, C are independent sets. This graph is clearly planar. Let $\mathscr{P}$ be a pairing of the vertices and let $\{u, v\} \in \mathscr{P}$. We describe how to join u and v by a path in all possible cases.

1. If there is an edge between u and v, join them by this edge.

2. If $u \in \{x_{AB}, x_{BC}, x_{CA}\}$ and $v \in A \cup B \cup C$ such that there is no edge between them, join them by the path uwv where the edge uw is consistent with the cyclic ordering x_{AB}, x_{BC}, x_{CA}. For example, if $u = x_{AB}$ and $v \in C$, we join u and v by the path $ux_{BC}v$. The remaining cases can be dealt using the same pattern.

3. If $u, v \in A \cup B \cup C$ and they are in the same class, join them by the path uwv where w is an arbitrary common neighbor (out of the two available).

4. If $u, v \in A \cup B \cup C$ and they are in different classes, join them by the path uwv where w is the unique common neighbor.

It is straightforward to check that the above instructions find edge-disjoint paths joining terminals, regardless of the choice of $\mathscr{P}$. □

1.2.2 The Proof of Theorem 1.2

The aim of this section is to prove our main theorem, Theorem 1.2. Our proof is based on three preparatory lemmas. First, we shall introduce some terminology. Let G be a multigraph. We say that two multiedges e, f of G are at distance d if the shortest path in G joining an endpoint of e and an endpoint of f has length d. If two multiedges are at distance 0, we shall simply say they are *incident*. Further, we shall refer to a matching of size k as a *k-matching*. We say that a k-matching is *good* if every pair of edges in the matching is at distance exactly 1. Notice that contracting all the edges of a good k-matching results in the complete graph K_k (with potential multiple edges and loops).

Our first lemma says that in any multigraph either some multiedges 'cluster' together or many pairs of multiedges are far apart, or one can find a good k-matching. We shall need the following inequality.

Fact 1.6. If $k \geq 2$ then $2^{-k}\left(\frac{1+2^{-k-1}}{(1-2^{-k})^2}\right) \leq 2^{-k+1}$.

The above inequality is easily seen to be equivalent to $(2^{-k+2}-1)(2^{-k-1}-1) \geq 0$.

Lemma 1.7. Let k be a natural number and $\varepsilon_1, \varepsilon_2$ be positive reals such that $\varepsilon_1 + \varepsilon_2 \leq 2^{-k}$. Then, for sufficiently large $M = M(k)$, if G is a multigraph on M multiedges, then at least one of the following conditions is satisfied.

1. There is a multiedge in G which is incident with at least $\varepsilon_1 M$ multiedges;
2. There are at least $\varepsilon_2 \binom{M}{2}$ pairs of multiedges which are at distance greater than 1;
3. G contains a good k-matching.

Proof. We shall use induction on k. The base case when $k = 1$ is trivial - Condition 3 is always satisfied. Assume then that $k \geq 2$ and the lemma is true for $k-1$.

Suppose every multiedge is incident with at most $\varepsilon_1 M$ multiedges and at most $\varepsilon_2 \binom{M}{2}$ pairs of multiedges are at distance greater than 1. We shall show that G contains a good k-matching. By an averaging argument there is a multiedge e which is at distance at most 1 from at least $(1-\varepsilon_2)M - 1$ multiedges. Let E' be the set of those multiedges which are at distance exactly 1 from e. It follows from our assumptions that
$M' = |E'| \geq (1-\varepsilon_1-\varepsilon_2)M - 1 \geq (1-2^{-k})M - 1$. Let G' be the multigraph spanned by E'. By assumption, at most $\varepsilon_2 \binom{M}{2}$ of the multiedges in G' are at distance greater than 1. Therefore, since $M \leq \frac{M'+1}{1-2^{-k}}$, for large enough M (and hence large enough M') we have that at most

$$\varepsilon_2 \binom{M}{2} \leq \varepsilon_2 \binom{\frac{M'+1}{1-2^{-k}}}{2} = \frac{\varepsilon_2}{(1-2^{-k})^2}\left(1+\frac{1}{M'}\right)\left(1+\frac{1+2^{-k}}{M'-1}\right)\binom{M'}{2}$$
$$\leq \frac{\varepsilon_2(1+2^{-k-1})}{(1-2^{-k})^2}\binom{M'}{2},$$

pairs of multiedges in G' are at distance greater than 1. Also, for M' large enough, each multiedge in G' is incident with at most
$\varepsilon_1 M \leq \varepsilon_1 \frac{M'+1}{1-2^{-k}} = \frac{\varepsilon_1}{1-2^{-k}}(1+\frac{1}{M'})M' \leq \frac{\varepsilon_1(1+2^{-k-1})}{1-2^{-k}} M'$ multiedges. Note that for $k \geq 2$ one

has

$$\varepsilon_1 \frac{1+2^{-k-1}}{1-2^{-k}} + \varepsilon_2 \frac{1+2^{-k-1}}{(1-2^{-k})^2} \leq \varepsilon_1 \frac{1+2^{-k-1}}{(1-2^{-k})^2} + \varepsilon_2 \frac{1+2^{-k-1}}{(1-2^{-k})^2}$$
$$\leq 2^{-k} \frac{1+2^{-k-1}}{(1-2^{-k})^2} \leq 2^{-(k-1)},$$

where the last inequality is precisely Fact 1.6. Therefore, by the induction hypothesis, G' contains a good $(k-1)$-matching. But since e is at distance 1 from any multiedge in G', we also have a good k-matching in G. □

Since we shall be operating with planar graphs, we single out the following corollary.

Corollary 1.8. Let M be a sufficiently large integer and let $\varepsilon_1, \varepsilon_2$ be positive reals such that $\varepsilon_1 + \varepsilon_2 \leq \frac{1}{32}$. If G is a planar multigraph with M multiedges then either G has a multiedge which is incident with at least $\varepsilon_1 M$ multiedges or there are at least $\varepsilon_2 \binom{M}{2}$ pairs of multiedges at distance greater than 1.

Proof. If G contained a good 5-matching then it would contain a K_5 minor. □

One strategy in the proof of our main theorem is to consider a suitable bipartition of our path-pairable planar graph, and to exploit the fact that any bipartite planar graph on n vertices has at most $2n-4$ edges. To exploit this last property we shall need ways of finding pairings of the vertices such that their corresponding edge-disjoint paths contribute 'many' edges to the bipartition. This is formalized in the following lemma.

Lemma 1.9. Let D be a positive integer and $0 < \varepsilon \leq 1/2$. Then there exists $c > 0$ such that the following is true. Suppose G is a path-pairable planar graph on $n > 1/c$ vertices with $\Delta = \Delta(G) \leq cn$. Let $A, U \subset V(G)$ be given with $U \subset A$ such that every vertex in A has degree at most D, $|A| \geq (1-\varepsilon)n$ and $|U| \geq \varepsilon n$. Let $B = V(G) \setminus A$. Then there is a pairing of the vertices in U which contributes to at least $2|U| - 16\varepsilon n$ edges between A and B.

Proof. We say that a path in G is *weak* if it begins and ends in A, uses no edges inside B, and uses at most 2 edges between A and B. Now, let $C := \lceil 4\varepsilon^{-1} \rceil$ and note that since

$\varepsilon \le 1/2$ we have that $\frac{3}{C-2} \le \varepsilon$. For every $x \in U$, let $U_x = \{u \in U : \exists \text{ weak } x-u \text{ path in } G \text{ of length at most } C\}$. We claim that U_x is small for every $x \in U$; namely, it is easy to see that

$$|U_x| \le D^C + D^C D\Delta D^C = D^C\left(1 + D^{C+1}\Delta\right).$$

Choose $c = c(D,\varepsilon) = \frac{\varepsilon}{4D^{2C+1}}$ so that $\Delta \le cn$. Then

$$|U_x| \le D^C\left(1 + D^{C+1}\Delta\right) \le \left(D^C + D^{2C+1}\right)\Delta \le 2D^{2C+1}\Delta \le 2D^{2C+1}cn \le \frac{\varepsilon}{2}n.$$

Let us define an auxiliary graph G_U with vertex set U where we join two vertices x, y provided $y \notin U_x$ (equivalently, $x \notin U_y$). It is easy to see that G_U has a perfect matching (or 'almost' perfect, if $|U|$ is odd; this makes no difference for us). Indeed, the degree of every vertex in G_U is at least $|U| - \frac{\varepsilon}{2}n \ge |U|/2$, and therefore G_U has a Hamilton cycle. Fix a perfect matching $\mathscr{M}$ in G_U according to this Hamilton cycle and fix a pairing $\mathscr{P}$ of the vertices of G where each edge of $\mathscr{M}$ forms a pair. Finally, since G is path-pairable, choose a collection of edge-disjoint paths $\mathscr{R}$ that realize this pairing. Observe that any path from $\mathscr{R}$ must use an even number of edges between A and B. We single out two types of edges $e = xy$ in $\mathscr{M}$ with respect to this realization: either the $x-y$ path in $\mathscr{R}$ is weak but is of length bigger than C, or this $x-y$ path uses at least 4 edges between A and B. Let $\mathscr{M} = \mathscr{E}_0 \cup \mathscr{E}_1 \cup \mathscr{E}_2$, where $\mathscr{E}_0$ denotes the edges satisfying the former condition, $\mathscr{E}_1$ the latter, and $\mathscr{E}_2$ denotes the remaining edges. We claim that most of the edges are in $\mathscr{E}_1$. Indeed, observe that if $e = xy \in \mathscr{E}_2$, then the $x-y$ path must use edges from B. By planarity we have $e(B) \le 3|B|$, and therefore $|\mathscr{E}_2| \le 3\varepsilon n$. Using planarity again we have that $e(A) \le 3|A|$. On the other hand, for each edge in $\mathscr{E}_0$ its path in $\mathscr{R}$ uses more than C edges, at most 2 of which are in the cut $\{A, B\}$, and none of which belong to B.

Accordingly, since these paths are edge-disjoint, we have that $e(A) \geq (C-2)|\mathscr{E}_0|$ and so

$$|\mathscr{E}_0| \leq \frac{3}{C-2}|A| \leq \varepsilon|A|.$$

Therefore, $|\mathscr{E}_1| \geq \frac{1}{2}(|U|-1) - |\mathscr{E}_0| - |\mathscr{E}_2| \geq \frac{1}{2}(|U|-1) - \varepsilon n - |\mathscr{E}_2|$. It follows that since every path in $\mathscr{R}$ pairing an edge in $\mathscr{E}_1$ contributes at least 4 edges between A and B, and these paths must be edge-disjoint, we have

$$e(A,B) \geq 4|\mathscr{E}_1| + 2|\mathscr{E}_2| \geq 2|U| - 2 - 4\varepsilon n - 2|\mathscr{E}_2| \geq 2|U| - 2 - 10\varepsilon n \geq 2|U| - 16\varepsilon n,$$

where in the last inequality we used the fact that $n > 1/c \geq 1/\varepsilon$. This completes the proof of Lemma 1.9. □

Our final lemma allows us to quantify more precisely the degree distribution in any bipartite planar graph.

Lemma 1.10. Let G be a bipartite planar graph on n vertices with parts A, B, and let $A' \subset A$ be the set of vertices in A with degree at least 3. Then the following are true.

1. The number of vertices in A with degree 2 is at least $e(A,B) - n - 3|B|$;
2. $|A'| < 2|B|$;
3. $e(A',B) < 6|B|$.

Proof. For each $i \geq 0$ let $A_i, A_{\leq i}$, and $A_{\geq i}$ denote the number of vertices in A that have degree i in G, degree at most i, and degree at least i, respectively. Because of planarity we have that $e(A',B) < 2(|A'| + |B|)$. Alternatively, $e(A',B) \geq 3|A'|$ so it follows that $A_{\geq 3} = |A'| < 2|B|$, and so $e(A',B) \leq 2(|A'| + |B|) < 6|B|$, establishing the second and third items. Further, we can bound the number of edges between A and B as

$$\begin{aligned} e(A,B) &\leq A_{\leq 1} + 2(|A| - A_{\leq 1} - A_{\geq 3}) + e(A',B) \\ &\leq A_{\leq 1} + 2(|A| - A_{\leq 1} - |A'|) + 2(|A'| + |B|) \\ &\leq 2|A| - A_{\leq 1} + 2|B|. \end{aligned}$$

It follows that $A_{\leq 1} \leq 2|A| + 2|B| - e(A,B)$. Finally, we see that
$A_2 = |A| - |A'| - A_{\leq 1} > e(A,B) - |A| - 4|B| = e(A,B) - n - 3|B|$, as required.

□

We are now in a position to prove our main theorem, Theorem 1.2. First, let us give a rough sketch of the proof. Let G be a path-pairable planar graph. We first partition the vertex set of G into the set A of vertices of small degree and the set B of vertices of large degree. We can apply Lemma 1.9 to find that there are many edges in this cut. We shall then show that most vertices in A have degree 2 in this bipartite graph. If $Y \subset A$ denotes the vertices of degree 2, then we define a planar multigraph with vertex set B where we join $x, y \in B$ whenever there is a $v \in Y$ joined to precisely x and y. Now, using Corollary 1.8, we are able to either find a vertex of linear degree in B, or we can find many pairs of multiedges in our multigraph that are far apart. This, however, allows us to find a pairing which contributes to more than $2n$ edges between A and B, a contradiction to planarity.

We restate Theorem 1.2 for convenience.

Theorem 1.2. There exists $c \geq 10^{-10^{10}}$ such that if G is a path-pairable planar graph on n vertices then $\Delta(G) \geq cn$.

Proof. Suppose G is a path-pairable planar graph and fix some large constant D so that $D^{-1} \leq 8.5 \cdot 10^{-6}$. Partition the vertex set of G into sets A and B, where $B = \{v \in V(G) : d(v) \geq D\}$ and $A = V(G) \setminus B$. Since $e(G) < 3n$ it easily follows that

$|B| \le 6D^{-1}n := \varepsilon n$. Suppose that $\Delta(G) < cn$, where c is sufficiently small (depending only on D) given by Lemma 1.9. More precisely, we may take

$$c = \frac{\varepsilon}{4D^{2\lceil 4/\varepsilon \rceil + 1}}.$$

Our aim is to obtain a contradiction to the planarity of G, and so there must exist a vertex of degree at least cn. Of course, this is trivial if $cn \le 1$, so we shall assume throughout that $n > 1/c$. By Lemma 1.9 (with $U = A$) we have that there are at least $2|A| - 16\varepsilon n \ge 2n - 18\varepsilon n$ edges between A and B.

Next, we shall show that there is a large subset of A which induces a graph with maximum degree at most 2. To see this, let $A_0 = A, B_0 = B$. Suppose A_i, B_i have been defined already. If there is a vertex $v \in A_i$ such that $d_{A_i}(v) > d_{B_i}(v)$, then let $A_{i+1} = A_i \setminus \{v\}$ and $B_{i+1} = B_i \cup \{v\}$. Notice that $e(A_{i+1}, B_{i+1}) \ge e(A_i, B_i) + 1$, and so $e(A_{i+1}, B_{i+1}) \ge e(A, B) + i \ge 2n - 18\varepsilon n + i$. Let $t \ge 0$ be such that there is no $v \in A_t$ with more neighbors in A_t than in B_t. Observe that $t \le 18\varepsilon n$ (otherwise $e(A_t, B_t) \ge 2n$), and accordingly $|B_t| = |B| + t \le \varepsilon n + 18\varepsilon n = 19\varepsilon n$.

Let $X \subset A_t$ be the set of vertices in A_t with at least 3 neighbors in A_t. Since every vertex in A_t has at least as many neighbors in B_t as in A_t, we have that every vertex in X has at least 3 neighbors in B_t. Therefore, by Lemma 1.10, $|X| \le 2|B_t|$, $e(X, B_t) \le 6|B_t|$, and there are at least $e(A_t, B_t) - n - 3|B_t| \ge e(A, B) - n - 3|B_t|$ vertices in A_t with exactly two neighbors in B_t. Let $A^* = A_t \setminus X$ and $B^* = B_t \cup X$. Now we have that every vertex in A^* has at most 2 neighbors in A^* and $|B^*| \le 3|B_t| \le 57\varepsilon n$, so $|A^*| \ge n - 57\varepsilon n$. We have to make sure we still have many vertices in A^* with exactly two neighbors in B^*. Notice that if a vertex $v \in A_t$ had two neighbors in B_t and was not adjacent to any vertex in X then $v \in A^*$ and v still has exactly two neighbors in B^*. Therefore we only have to worry about the vertices in A_t which are adjacent to some vertices in X. Observe that $e(X, A^*) \le e(X, B_t) \le 6|B_t|$, and so there are at least

$e(A,B) - n - 9|B_t| \geq (2n - 18\varepsilon n) - n - 9 \cdot 19\varepsilon n = n - 189\varepsilon n$ vertices in A^* with exactly 2 neighbors in B^*. Hence there are at most $189\varepsilon n$ vertices in A^* which do not have degree 2 in B^*.

We say that an edge $uv \in G$ is *bad* if one of the followings holds:

1. (Type I) $uv \in G[B^*]$.

2. (Type II) $uv \in G[A^*]$ and u (or v) has degree not equal to 2 in B^*.

3. (Type III) $uv \in G[A^*]$, $d_{B*}(u) = d_{B*}(v) = 2$, and $N_{B^*}(u) \neq N_{B^*}(v)$.

4. (Type IV) $uv \in G$, such that $u \in A^*, v \in B^*$, and $d_{B^*}(u) \geq 3$.

We have the following bound on the number of bad edges.

Claim 1.11. There are at most $1233\varepsilon n$ bad edges.

Proof. We are going to bound the number of bad edges of each type.

Note that by planarity, there are at most $3|B^*|$ edges in B^* so there are at most $3|B^*| \leq 171\varepsilon n$ edges of Type I.

Now, since every vertex in A^* has at most two neighbors in A^*, each vertex in A^* with degree not equal to 2 in B^* contributes to at most two bad edges of Type II. As there are at most $189\varepsilon n$ vertices in A^* which do not have degree 2 in B^*, it follows that there are at most $378\varepsilon n$ bad edges of Type II.

Let us consider bad edges of Type III. Since $G[A^*]$ has maximum degree 2, we can partition the edges of $G[A^*]$ into at most 3 matchings, M_1, M_2, M_3. It is well known (and easy to see) that contracting an edge in a planar graph preserves planarity. It follows that, for $i \in \{1,2,3\}$, we can contract the edges of M_i while still preserving planarity. Denote this new graph by $\tilde{G}_i$ with vertex set $\tilde{A}_i \cup B^*$. Since $\tilde{G}_i$ is planar, from Lemma 1.10 we have that there are at most $2|B^*|$ vertices in $\tilde{G}_i$ with at least 3 neighbors in B^*. Therefore, at most $2|B^*|$ edges in M_i can be bad of Type III. Hence, there are at most $6|B^*| \leq 342\varepsilon n$ bad edges of Type III.

Finally, by Lemma 1.10 there can be at most $6|B^*| \leq 342\varepsilon n$ bad edges of Type IV.

So in total there are at most $1233\varepsilon n$ bad edges of any type. □

Let $Y \subseteq A^*$ be the set of vertices with degree exactly 2 in B^*. We now define an auxiliary *multi*graph G_{B^*} in the following way. The vertex set of G_{B^*} is B^* and for any two vertices $x, y \in B^*$, join x to y by an edge for every $v \in Y$ that is joined precisely to x and y.

Claim 1.12. G_{B^*} is planar.

Proof. This is clear since the bipartite graph $G[Y, B^*]$ between Y and B^* is planar, and contracting edges preserves planarity. □

Apply Corollary 1.8 to the multigraph G_{B^*} with $\varepsilon_1 = \varepsilon_2 = 1/100$. Notice that if an edge in G_{B^*} is incident to more than $\frac{1}{100}|Y|$ multiedges then one of its endpoints has degree, in G, at least $\frac{1}{200}|Y|$. However, recall that we initially assumed $\Delta(G) < cn$, and certainly $c \leq 1/400$ by our choice of D. Accordingly, since $|Y| \geq n - 189\varepsilon n \geq n/2$, we obtain a vertex of degree at least

$$2c|Y| \geq cn,$$

a contradiction.

So we may assume that there are at least $\frac{1}{100}\binom{|Y|}{2}$ pairs of edges in G_{B^*} which are at distance greater than 1. Note that if H is any graph on n vertices with edge density at least δ, then it is easy to greedily find a matching of size at least $\frac{\delta}{10}n$. It follows that we may select a collection of pairwise disjoint pairs $\mathscr{P}$ in Y, such that $|\mathscr{P}| \geq \frac{1}{1000}|Y| \geq \frac{1}{2000}n$, and such that for every $\{u, v\} \in \mathscr{P}$, their corresponding edges in G_{B^*} are at distance greater than 1.

We need the following two claims.

Claim 1.13. Let P be a path contained in A^* which has at least two vertices and does not contain any bad edges. Then every vertex $v \in P$ has the same neighborhood (of size 2) in B^*.

Proof. This is immediate from the definition of a bad edge. □

Claim 1.14. Let $u, v \in Y$ be two vertices whose corresponding edges in G_{B^*} are at distance greater than 1. Then any path in G joining u and v either contains some bad edges, or uses at least 6 edges between A^* and B^*.

Proof. Suppose P is a path joining u and v which does not use any bad edges and does not use at least 6 edges between A^* and B^*. By definition and using claim 1.13, all vertices of $V(P) \cap A^*$ are in Y, it can not have an edge inside B^* and it must use 2 or 4 edges between A^* and B^*. We may assume P uses 4 edges as the other case follows from the same argument. Let $P = P_1 e_1 e_2 P_2 e_3 e_4 P_3$, where $\{e_1, e_2, e_3, e_4\}$ are edges between A^* and B^* and P_1, P_2, P_3 are paths inside Y. From claim 1.13 applied to P_1, P_2 and P_3 we deduce that the edge of u in G_{B^*} is at distance at most 1 to the edge of v in G_{B^*}. □

The proof of Theorem 1.2 is nearly complete. Indeed, since G is path-pairable, there are edge-disjoint paths joining every pair of $\mathscr{P}$, and hence the pairs in $\mathscr{P}$ contribute to at least $6(|\mathscr{P}| - 1233\varepsilon n)$ edges between A^* and B^*.

Let P be the union of the vertices in $\mathscr{P}$ and let $U = A^* \setminus P$. Suppose first that $|U| < 57\varepsilon n$. It follows that

$$2|\mathscr{P}| > (n - 57\varepsilon n) - 57\varepsilon n,$$

so $|\mathscr{P}| > n/2 - 57\varepsilon n$. Then the above pairing contributes at least $6(n/2 - 1290\varepsilon n) = 3n - 7740\varepsilon n$ edges between A^* and B^*. But this is at least $2n$ whenever $\varepsilon \leq 7740^{-1}$ which is guaranteed by our choice of D, a contradiction. Therefore, we may assume that $|U| \geq 57\varepsilon n$. By Lemma 1.9 (since c is small enough) there is a pairing of the vertices in U which contributes to at least $2|U| - 16 \cdot 57\varepsilon n = 2|U| - 912\varepsilon n$

edges between A^* and B^*. Hence in total the number of edges between A^* and B^* is

$$\begin{aligned} &\geq 6(|\mathscr{P}| - 1233\varepsilon n) + 2|A^*| - 4|\mathscr{P}| - 912\varepsilon n \\ &\geq 2|\mathscr{P}| + 2(n - 57\varepsilon n) - 6 \cdot 1233\varepsilon n - 912\varepsilon n \\ &\geq 2n + n/1000 - 8424\varepsilon n. \end{aligned}$$

So by our choice of D we get that $8424\varepsilon \leq \frac{1}{1000}$, and so there are at least $2n$ edges between A^* and B^*, a contradiction to the planarity of G. It follows that there must exist a vertex of degree at least cn.

□

1.2.3 Final Remarks and Open Problems

It is worth observing that our proof relies only on the following three properties of a planar graph G: contracting edges of G preserves planarity, G does not contain a K_5-minor, and any bipartite subgraph H of G has at most $2|H|$ edges. We remark that it is possible to generalize our result in the following sense. Given integers t, c, we say that a graph G is (t,c)-*good* if G is K_t-minor-free and any bipartite subgraph H of G has at most $2|H| + c$ edges. Moreover, define $\mathscr{G}_{t,c}$ to be the family of (t,c)-good graphs G such that contracting edges of G preserves (t,c)-goodness.

Theorem 1.15. For any integers t, c there is a positive constant $C = C(t,c)$ such that the following holds. If G is a path-pairable graph on n vertices with $G \in \mathscr{G}_{t,c}$, then $\Delta(G) \geq Cn$.

We have the following immediate corollary.

Corollary 1.16. For every non-negative integer g there is a positive constant $C = C(g)$ such that the following holds. If G is a path-pairable graph on n vertices which has a 2-cell embedding on a surface with genus g, then $\Delta(G) \geq Cn$.

Proof. We claim that $G \in \mathscr{G}_{3g+5,2g}$. Indeed, it follows from Euler's formula (see, e.g., [13]) that if G is 2-cell embedded on a surface of genus g then $n + m - f = 2 - g$, where m

is the number of edges G and f is the number of faces of the embedding. Since $2m \geq 3f$ ($2m \geq 4f$ if G is triangle-free) it follows that $e(G) \leq 3n+3g-6$ ($e(G) \leq 2n+2g-4$ if G is triangle-free). In particular, if G is bipartite then $e(G) \leq 2n+2g-4 \leq 2n+2g$. Suppose for contradiction that G contains a K_{3g+5}-minor. Then K_{3g+5} could be 2-cell embedded on a surface of genus g, hence $\binom{3g+5}{2} = e(K_{3g+5}) \leq 12g+9$, which is easily seen to be a contradiction. □

Sketch of a proof of Theorem 1.15. The proof is essentially the same as the proof of Theorem 1.2. Certain changes have to be made in the preparatory lemmas first.

Corollary 1.8 generalizes trivially to multigraphs with no K_t-minors.

In Lemma 1.9 we only use the fact that any subgraph H of a planar graph has at most $3|H|$ edges. Observe that if $G \in \mathscr{G}_{t,c}$ then any subgraph of H of G has at most $4|H|+2c$ edges. One can therefore modify the proof, at the expense of a worse constant in front of εn in the conclusion of the Lemma.

In the proof of Lemma 1.10 we only use the fact that a bipartite subgraph H of a planar graph does not use more than $2|H|$ edges. The lemma can be therefore modified to work for graphs in $\mathscr{G}_{t,c}$ by introducing some additive constants, depending only on c, to the inequalities in every part of the Lemma.

In the proof of Theorem 1.2 all the estimates remain correct by taking ε small enough.

Note that we also need that $\mathscr{G}_{t,c}$ is closed under edge contractions in order to estimate the number of "bad edges", as in Claim 1.11 (there we used that contracting edges preserves planarity). □

We believe that the condition on the number of edges in bipartite subgraphs can be omitted while still ensuring the existence vertex of linear degree. We therefore make the following conjecture.

Conjecture 1.17. For any t there exists a constant $c = c(t)$ such that every path-pairable graph on n vertices without a K_t minor must contain a vertex of degree at least cn.

Finally, recall that we defined $\Delta^p_{\min}(n)$ to be the minimum of $\Delta(G)$ over all n-vertex path-pairable planar graphs G. We have shown that $\Delta^p_{\min}(n)$ grows linearly in n; however, as mentioned in the Introduction, the constants in the upper and lower bounds are quite far apart. We close with the following problem.

Problem 1.18. Determine $\Delta^p_{\min}(n)$ for sufficiently large n.

We do not know if our construction yielding the upper bound of $2n/3$ is optimal, and a significant improvement on our lower bound would be very interesting.

1.3 Diameter of path-pairable graphs

1.3.1 Path-pairable graphs from blowing up paths

In this section, we shall show how to construct a class of graphs which have large diameter and are path-pairable. Let G be a graph with vertex set $V(G) = \{v_1, \dots, v_k\}$, and let $G_1, \dots, G_k$ be graphs. We define the *blown-up graph* $G(G_1, \dots, G_k)$ as follows: replace every vertex v_i in G by the corresponding graph G_i, and for every edge $v_iv_j \in E(G)$ insert a complete bipartite graph between the vertex sets of G_i and G_j.

Let P_k denote the path on k vertices. The following lemma asserts that if we blow up a path with graphs $G_1, \dots, G_k$, such that G_i is path-pairable for $i \leq k-1$, and certain properties inherited from the cut-condition hold, then the resulting blow-up is path-pairable.

Lemma 1.19. Let $G_1, \dots, G_k$ be graphs on $n_1, \dots, n_k$ vertices, respectively, where G_i is path-pairable for $i \leq k-1$. Let $n = \sum_{i=1}^k n_i$ and let $u_i = \sum_{j=1}^i n_j$ for $i = 1, \dots, k-1$. Then $P_k(G_1, \dots, G_k)$ is path-pairable if and only if

$$n_i \cdot n_{i+1} \geq \min(u_i, n - u_i) \tag{1.1}$$

holds for $i = 1, \ldots, k-1$.

Proof. For each $i = 1, \ldots, k$, let $U_i = \bigcup_{j=1}^{i} V(G_j)$ so that $u_i = |U_i|$. Now, if $P_k(G_1, \ldots, G_k)$ is path-pairable, then we may apply the cut-condition to the cut $\{U_i, V(G) \setminus U_i\}$. This implies $n_i \cdot n_{i+1} \geq \min(u_i, n - u_i)$ must hold for $i = 1, \ldots, k-1$. In the remainder, we show that this simple condition is enough to yield the path-pairability of $G := P_k(G_1, \ldots, G_k)$. Assume that a pairing $\mathscr{P}$ of the vertices of G is given. If $\{u, v\} \in \mathscr{P}$ we shall say that u is a *sibling* of v (and vice-versa). We shall define an algorithm that sweeps through the classes $G_1, G_2, \ldots, G_k$ and joins each pair of siblings via edge-disjoint paths.

First we give an overview of the algorithm. We proceed by first joining pairs $\{u, v\} \in \mathscr{P}$ via edge-disjoint paths such that u and v belong to different G_i's, and then afterwards joining pairs that remain inside some G_j (using the path-pairability of G_j). Before round 1, we use the path-pairability property of G_1 to join those siblings which belong to G_1. During round 1, we assign to every vertex u of G_1 a vertex v of G_2. If $\{u, v\} \in \mathscr{P}$ are siblings, then we simply choose the edge uv. Then we join the siblings which are in G_2 again using the path-pairability property of G_2. For those paths uv that have not ended (because $\{u, v\} \notin \mathscr{P}$) we shall continue by choosing a new vertex w in G_3 and continue the path with edge vw, and so on. We shall call paths which have not finished joining a pair of siblings *unfinished*; otherwise, we say the path is *finished*. The last edge which completes a finished path we shall call a *path-ending edge*. During round i, we shall first choose those vertices in G_{i+1} which, together with some vertex of G_i, form path-ending edges. At the end of round i, in G_{i+1} we will have endpoints of unfinished paths and perhaps also some endpoints of finished paths. Note that the vertices of G_{i+1} might be endpoints of several unfinished paths. For $x \in G_{i+1}$ let $w(x)$ denote the number of unfinished paths $P \cup \{x\}$ with $P \subset U_i$ at the end of round i which are to be extended by a vertex of G_{i+2} (including the single-vertex path x in the case when x was not joined to its sibling in the latest round). Note that every such path corresponds to a yet not joined vertex in U_{i+1} as well as to another vertex yet to be joined lying in $V(G) \setminus U_{i+1}$. It follows that

$$\sum_{x \in G_{i+1}} w(x) \leq \min(u_{i+1}, n - u_{i+1}). \tag{1.2}$$

Let us now be more explicit in how we make choices in each round. We shall maintain the following two simple conditions throughout our procedure (the first of which has been mentioned above):

(a) During round i ($1 \leq i \leq k-1$), if $w \in G_i$ is the current endpoint of the path which began at some vertex $u \in U_i$ (possibly $u = w$), and $\{u, v\} \in \mathscr{P}$ for $v \in G_{i+1}$, then we join w to v. Informally, we choose path-ending edges when we can.

(b) $w(x) \leq n_{i+1}$ for all $x \in G_i$, for $i = 1, \ldots, k-1$.

The second condition above is clearly necessary in order to proceed during round i, as $|N(x) \cap G_{i+1}| = n_{i+1}$ for every $x \in G_i$, and hence we cannot continue more than n_{i+1} unfinished paths through x.

We claim that as long as both of the above conditions are maintained, the proposed algorithm finds a collection of edge-disjoint paths joining every pair in $\mathscr{P}$. Both conditions are clearly satisfied for $i = 1$ as $w(x) \leq 1 \leq n_2$ for all $x \in G_1$. Let $i \geq 2$ and suppose both conditions hold for rounds $1, \ldots, i-1$. Our aim is to show that an appropriate selection of edges between G_i and G_{i+1} exists in round i to maintain the conditions. We start round i by choosing all path-ending edges with endpoints in G_i and G_{i+1}; this can be done since, by induction, $w(x) \leq n_{i+1}$ for every $x \in G_i$. Observe that if $i = k-1$ then the only remaining siblings are in G_k. Then for every $\{u, v\} \in \mathscr{P}$ such that $u, v \in G_k$ we can find a vertex q in G_{k-1} and join u, v with the path uqv. When $i < k-1$ then the remaining paths can be continued by assigning arbitrary vertices from G_{i+1} (without using any edge multiple times). We choose an assignment that balances the 'weights' in G_{i+1}. More precisely, let us choose an assignment of the vertices that minimizes

$$\sum_{a \in G_{i+1}} (w(a))^2.$$

If for every $x \in G_{i+1}$ we have that $w(x) \leq n_{i+2}$ we are basically done. It remains to find edge-disjoint paths inside G_{i+1} for those pairs $\{x, y\} \in \mathscr{P}$ whose vertices belong to G_{i+1}. But this is possible because of the assumption that G_{i+1} is path-pairable.

Suppose then that in the above assignment there exists $x \in G_{i+1}$ with $w(x) \geq n_{i+2} + 1$. We first claim that, under this assignment, no other vertex of G_{i+1} has small weight.

Claim 1.20. Every vertex $y \in G_{i+1}$ satisfies $w(y) \geq n_{i+2} - 1$.

Proof. Suppose there is $y \in G_{i+1}$ such that $w(y) \leq n_{i+2} - 2$. Then, as $w(x) > w(y) + 2$, there exist vertices $v_1, v_2 \in G_i$ such that a nonempty collection of paths ending at v_1 and v_2 have x as their next vertex after round i, but no paths using v_1 or v_2 are assigned y as their next vertex. Note that at least one of the edges v_1x, v_2x is not path-ending; we may assume v_1x is not path-ending. Replacing our choice of v_1x with v_1y decreases the square sum $\sum_{a \in G_{i+1}} (w(a))^2$—a contradiction. □

Therefore, we may assume $w(y) \geq n_{i+2} - 1$ for all $y \in G_{i+1}$. In this case, partition the vertices of G_{i+1} into three classes:

$$X = \{v \in G_{i+1} : w(v) \geq n_{i+2} + 1\}$$
$$Y = \{v \in G_{i+1} : w(v) = n_{i+2} - 1\}$$
$$Z = \{v \in G_{i+1} : w(v) = n_{i+2}\}.$$

Observe first that $1 \leq |X| \leq |Y|$, since otherwise, using (1.2), we have

$$n_{i+1}n_{i+2} + 1 \leq \sum_{s \in G_{i+1}} w(s) \leq \min(u_{i+1}, n - u_{i+1}),$$

contradicting condition (1.1). Notice also that a similar argument as in Claim 1.20 shows that $w(v) \leq n_{i+2}+1$ for every $v \in G_{i+1}$. Hence, we actually have

$$X = \{v \in G_{i+1} : w(v) = n_{i+2}+1\}.$$

We will need the following claim, which asserts that if there are siblings in G_{i+1}, then they must belong to Z.

Claim 1.21. If $\{u,v\} \in \mathscr{P}$ and $u,v \in G_{i+1}$, then $u,v \in Z$.

Proof. We first show that every $y \in Y$ is incident to a path-ending edge. Suppose, to the contrary, that there is $y \in Y$ with no path-ending edge ending at y. It follows that at most $w(y)$ vertices in G_i are joined to y after round i. Hence, we can take any $x \in X$ and find $z \in G_i$ such that z is not joined to y after round i, and such that zx is not path-ending. Replacing zx by zy results in a smaller square sum $\sum_{a \in G_{i+1}} (w(a))^2$, which gives a contradiction.

Now, let $\{u,v\} \in \mathscr{P}$ such that $u,v \in G_{i+1}$. Since every $y \in Y$ is incident to a path-ending edge, we have that $u,v \notin Y$. Suppose, for contradiction, that $u \in X$. Then u is joined to $w(u) = n_{i+2}+1$ vertices in G_i after round i, and so for every $y \in Y$ there is $z \in G_i$ which is joined to u but not to y. Replacing zu by zy results in a smaller square sum $\sum_{a \in G_{i+1}} (w(a))^2$, which again is a contradiction. □

Finally, we shall show how to reduce the weights of vertices in X (and pair the siblings inside G_{i+1}) using the path-pairable property of G_{i+1}. For every $x \in X$ pick a different vertex $y_x \in Y$ (which we can do, since $|Y| \geq |X|$) and let $\mathscr{P}' = \{\{u,v\} \in \mathscr{P} : u,v \in G_{i+1}\} \cup \{\{x,y_x\} : x \in X\}$. Since G_{i+1} is path-pairable, we can find edge-disjoint paths joining the siblings in $\mathscr{P}'$ (note that by Claim 1.21 none of the pairs $\{x,y_x\}$ interfere with any siblings $\{u,v\} \in \mathscr{P}$ with $u,v \in G_{i+1}$). Observe now that for every $x \in X$ one path has been channeled to a vertex $y \in Y$, and so the number of unfinished path endpoints at x has dropped to n_{i+2}, as required. This completes the proof of Lemma 1.19.

□

We close this section by pointing out that the path-pairability of certain G_i subgraphs in a path-pairable graph $P_k(G_1,\dots,G_k)$ cannot be avoided for $k \geq 5$. We do this by giving an example of a blown-up path that satisfies the cut-conditions of Lemma 1.19, but is not path-pairable unless some of the G_i's are path-pairable. For the sake of simplicity we set $k=5$ and prove that G_3 has to be path-pairable. Let $n = 2t^2+t$ for some even $t \in \mathbb{N}$, and let $n_1 = n_5 = t^2 - t$, $n_2 = n_3 = n_4 = t$, where $n_i = |G_i|$ for each $i \in [5]$. Clearly, $P_5(G_1,\dots,G_5)$ satisfies the condition (1.1) of Lemma 1.19. Also, observe that any pairing of the vertices in $G_1 \cup G_2$ with the vertices in $G_4 \cup G_5$ has to use all edges between G_3 and $G_2 \cup G_4$. Therefore, if we additionally pair the vertices inside G_3, then the paths joining those vertices can only use edges in G_3. Accordingly, G_3 must be path-pairable.

1.3.2 Proof of Theorem 1.4

Take $x,y \in V(G)$ such that $d(x,y) = d(G)$ and let V_i be the set of vertices at distance exactly i from x, for every i. Observe that $V_0 = \{x\}$ and $y \in V_{d(G)}$. For $i \in \{1,\dots,d(G)\}$ define n_i to be the size of V_i and let $u_i = \sum_{j=0}^{i} n_j$.

We need the following claim.

Claim 1.22. $u_{2k+1} \geq \binom{k+2}{2}$ as long as $u_{2k+1} \leq \frac{n}{2}$.

Proof. We shall use induction on k. For $k=0$ the assertion is clear. Assume that $u_{2k-1} \geq \binom{k+1}{2}$. By the cut-condition we have that the number of edges between V_{2k} and V_{2k+1} is at least u_{2k-1}, hence $n_{2k} \cdot n_{2k+1} \geq u_{2k-1} \geq \binom{k+1}{2}$. By the arithmetic-geometric mean inequality, $n_{2k} + n_{2k+1} \geq 2\sqrt{\binom{k+1}{2}} \geq k+1$. As $u_{2k+1} = u_{2k-1} + n_{2k} + n_{2k+1}$, we have $u_{2k+1} \geq \binom{k+2}{2}$. □

Now, let $A = \bigcup_{i=0}^{\lfloor d/3 \rfloor} V_i$, $B = \bigcup_{i=\lfloor d/3 \rfloor+1}^{\lfloor 2d/3 \rfloor} V_i$, $C = \bigcup_{i=\lfloor 2d/3 \rfloor+1}^{d} V_i$. Observe that $|A|,|C| \geq \min\left\{\frac{n}{2}, \frac{d^2}{100}\right\}$, so joining vertices in A with vertices in C requires at least

$\min\left\{\frac{n}{2}, \frac{d^2}{100}\right\} \cdot \frac{d}{3}$ edges. Hence,

$$\min\left\{\frac{n}{2}, \frac{d^2}{100}\right\} \cdot \frac{d}{3} \leq m,$$

which implies

$$d \leq \max\left\{\frac{6m}{n}, \sqrt[3]{300m}\right\}.$$

Notice that whenever $m \leq n^{3/2}$ we have $d \leq \sqrt[3]{300m}$. Also, if $m \geq \sqrt{2}n^{3/2}$, then the upper bound is trivially satisfied by the general upper bound obtained in [65].

For the lower bound, let n and $2n \leq m \leq \frac{1}{4}n^{3/2}$ be given. For any natural number ℓ we shall denote by S_ℓ the star $K_{1,\ell-1}$ on ℓ vertices. Consider the graph

$$G = P_{k+3}(G_1, \dots, G_{k+3})$$

on n vertices, where $k = \left\lfloor \sqrt[3]{\frac{m}{2} - n} \right\rfloor$ and $G_1 = G_2 = \cdots = G_k = S_k$, $G_{k+1} = S_{k^2}$, $G_{k+2} = S_2$, and G_{k+3} is an empty graph on $n - 2k^2 - 2$ vertices.

Straightforward calculation shows that

- $u_i = i \cdot k$, for $i \leq k$, $\quad u_{k+1} = 2k^2$, $\quad u_{k+2} = 2k^2 + 2$.
- $n_1 n_2 = n_2 n_3 = \cdots = n_{k-1} n_k = k^2$, $\; n_k n_{k+1} = k^3$, $\quad n_{k+1} n_{k+2} = 2k^2$, $n_{k+2} n_{k+3} = 2n - 4k^2 - 4$.

Therefore, for $i \in \{1, \dots, k+1\}$ we have

$$n_i \cdot n_{i+1} \geq u_i \geq \min(u_i, n - u_i),$$

and

$$n_{k+2} \cdot n_{k+3} \geq n_{k+3} \geq \min(u_{k+2}, n - u_{k+2}).$$

Hence, it follows from Lemma 1.19 that G is path-pairable.

It is easy to check that the number of edges in G is at most $2n+2k^3 \leq m$. On the other hand, the diameter of G is $k+2 \geq \sqrt[3]{\frac{m}{2}-n}$. This completes the proof of Theorem 1.4.

1.3.3 Proof of Theorem 1.5

In this section, we investigate the maximum diameter a path-pairable c-degenerate graph on n vertices can have. We shall assume that c is an integer and $c \geq 5$.

Let G be a c-degenerate graph on n vertices with diameter d. We shall show first that $d \leq 4\log_{\frac{c+1}{c}}(n)+3$. Let $x \in G$ be such that there is $y \in G$ with $d(x,y) = d$. For $i \in \{0,\dots,d\}$, write V_i for the set of vertices at distance i from x. Let $n_i = |V_i|$ and $u_i = \sum_{j=0}^{i} n_j$. Observe that $|V_i| \geq 1$ for every $i \in \{0,\dots,d\}$. Moreover, we can assume that $u_{\lfloor \frac{d}{2} \rfloor} \leq \frac{n}{2}$ (otherwise, repeat the argument below with $V_i' = V_{d-i}$).

The claimed upper bound on the diameter easily follows from the following claim.

Claim 1.23. $u_{2k+1} \geq \left(\frac{c+1}{c}\right)^k$ as long as $u_{2k+1} \leq \frac{n}{2}$.

Let us assume the claim and prove the result. Letting $k = \frac{\lfloor \frac{d}{2} \rfloor - 1}{2}$, we have that

$$\left(\frac{c+1}{c}\right)^k \leq u_{2k+1} \leq n/2,$$

by the above claim. Hence,

$$\begin{aligned} d &\leq 4\log_{\frac{c+1}{c}}(n)+3 = \frac{4\log(n)}{\log(\frac{c+1}{c})}+3 \leq \frac{4\log(n)}{\log(\frac{c+1}{c-1})}\frac{\log(\frac{c+1}{c-1})}{\log(\frac{c+1}{c})}+3 \\ &\leq 12\log_{\frac{c+1}{c-1}}(n)+3, \end{aligned}$$

where the last inequality follows from the easy to check fact that $\frac{\log(\frac{c+1}{c-1})}{\log(\frac{c+1}{c})} \leq 3$, for all $c \geq 5$. Thus it remains to prove the claim.

Proof of the Claim. We shall prove the claim by induction on k. The base case when $k = 0$ is trivial as $u_1 \geq 2$. Suppose the claim holds for every $l \leq k-1$. Since G is c-degenerate, we have that $e(V_{2k}, V_{2k+1}) \le c\,(n_{2k} + n_{2k+1})$. On the other hand, it follows from the cut-condition that $e(V_{2k}, V_{2k+1}) \ge u_{2k} = u_{2k-1} + n_{2k}$. Therefore, by the induction hypothesis, we have

$$\begin{aligned} n_{2k} + n_{2k+1} \ge \frac{1}{c}\,(u_{2k-1} + n_{2k}) &\ge \frac{1}{c}\left(\left(\frac{c+1}{c}\right)^{k-1} + n_{2k}\right) \\ &\ge \frac{1}{c}\left(\frac{c+1}{c}\right)^{k-1}. \end{aligned}$$

It follows that,

$$u_{2k+1} = u_{2k-1} + n_{2k} + n_{2k+1} \ge \left(\frac{c+1}{c}\right)^{k-1} + \frac{1}{c}\left(\frac{c+1}{c}\right)^{k-1}.$$

But the right-hand side is equal to $\left(1+\frac{1}{c}\right)\left(\frac{c+1}{c}\right)^{k-1} = \left(\frac{c+1}{c}\right)^{k}$, which proves the claim. □

We shall prove the lower bound in Theorem 1.5 assuming that c is an odd integer; when c is even we apply the same argument for $c-1$.

To do so, consider the graph $G = P_{2m'-1}(G_1, \ldots, G_{2m'-1})$ for some $m' \in \mathbb{N}$, which we specify later. First, we shall define the sizes of G_i for $i \in \{1, \ldots, 2m'-1\}$. To do so, define a sequence $\{n_i\}_{i\in\mathbb{N}}$ where $n_1 = 1$, $n_{2i} = \frac{c-1}{2}$ and n_{2i+1} is defined recursively in the following way:

$$n_{2i+1} = \left\lceil \frac{2}{c-1} \cdot \sum_{j=1}^{2i} n_j \right\rceil \le \left\lceil \frac{2}{c-1} \sum_{j=1}^{2i-2} \right\rceil + \left\lceil \frac{2}{c-1}(n_{2i-1} + n_{2i}) \right\rceil \tag{1.3}$$

$$\le n_{2i-1} + \left\lceil \frac{2}{c-1} n_{2i-1} + 1 \right\rceil \le \frac{c+1}{c-1} n_{2i-1} + 2 \le c\left(\frac{c+1}{c-1}\right)^{i} - (c-1), \tag{1.4}$$

where the last inequality can be easily proved by induction.

Let m be the largest integer such that $\sum_{j=1}^{m} n_j \le n/2$. Let $m' = m$ when m is odd, and $m' = m-1$ when m is even. Moreover, let $|G_{m'}| = n - 2\sum_{j=1}^{m'-1} n_j$ and let $|G_i| = n_i$ for

$1 \le i < m'$ and $|G_{m'+j}| = |G_{m'-j}|$ for $j \in \{1,\dots,m'-1\}$.

For all $i \in \{1,\dots 2m'-1\}$, let $G_i = S_{n_i}$ be a star on n_i vertices. It is easy to check that the graph $P_{2m'-1}(G_1,\dots,G_{2m'-1})$ is path-pairable by Lemma 1.19. It has diameter at least $2m-4$ and $m \ge 2\log_{\frac{c+1}{c-1}}(n) - \Theta_c(1)$, which follows from (1.3). Indeed, note that for $i \ge 1$,

$$\sum_{j=1}^{2i+1} n_j \le \frac{c-1}{2}\left[\frac{2}{c-1}\sum_{j=1}^{2i} n_j\right] + n_{2i+1} = \frac{c-1}{2}n_{2i+1} + n_{2i+1} = \frac{c+1}{2}n_{2i+1} \le \frac{c(c+1)}{2}\left(\frac{c+1}{c-1}\right)^i.$$

Lastly, it is not too hard to see that G is c-degenerate. Indeed, consider some ordering $v_1,\dots,v_n$ of the vertices of G such that if $v_i \in G_l$ and $v_{i'} \in G_{l'}$ (where l is even and l' is odd), then $i < i'$; and if $v_i, v_{i'} \in G_l$ and v_i is the center of the star G_l, then $i < i'$. In such an ordering, any vertex v_i is adjacent to at most $2 \cdot \frac{c-1}{2} + 1 = c$ vertices v_j with $j < i$. This proves that G is c-degenerate, and completes the proof of Theorem 1.5.

1.3.4 Final remarks and open problems

We obtained tight bounds on the parameter $d(n,\mathscr{G}_m)$ when $(2+\varepsilon)n \le m \le \frac{1}{4}n^{3/2}$, for any fixed $\varepsilon > 0$. It is an interesting open problem to investigate what happens when the number of edges in a path-pairable graph on n vertices is around $2n$. We ask the following:

Question 1.24. Is there a function f such that for every $\varepsilon > 0$ and for every path-pairable graph G on n vertices with at most $(2-\varepsilon)n$ edges, the diameter of G is bounded by $f(\varepsilon)$?

Another line of research concerns determining the behavior of $d(n,\mathscr{P})$, where $\mathscr{P}$ is the family of planar graphs. Since planar graphs are 5-degenerate, it follows from Theorem 1.5 that the diameter of a path-pairable planar graph on n vertices cannot be larger than $c\log n$. This fact makes us wonder whether there are path-pairable planar graphs with unbounded diameter.

Question 1.25. Is there a family of path-pairable planar graphs with arbitrarily large diameter?

The graph constructed in the proof of the lower bound in Theorem 1.5 when $c = 5$ is not planar since it contains a copy of $K_{3,3}$. Therefore, it cannot be used to show that the diameter of a path-pairable planar graph can be arbitrarily large (note, however, that this graph does not contain a K_7-minor nor a $K_{6,6}$-minor). We end by remarking that we were able to construct an infinite family of path-pairable planar graphs with diameter 6, but not larger.

CHAPTER 2

PARTITE SATURATION OF COMPLETE GRAPHS

In this chapter we study the problem of determining $sat(n,k,r)$, the minimum number of edges in a k-partite graph G with n vertices in each part such that G is K_r-free but the addition of an edge joining any two non-adjacent vertices from different parts creates a K_r. Improving recent results of Ferrara, Jacobson, Pfender and Wenger, and generalizing a recent result of Roberts, we define a function $\alpha(k,r)$ such that $sat(n,k,r) = \alpha(k,r)n + o(n)$ as $n \to \infty$. Moreover, we prove that

$$k(2r-4) \leq \alpha(k,r) \leq \begin{cases} (k-1)(4r-k-6) & \text{for } r \leq k \leq 2r-3, \\ (k-1)(2r-3) & \text{for } k \geq 2r-3, \end{cases}$$

and show that the lower bound is tight for infinitely many values of r and every $k \geq 2r-1$. This allows us to prove that, for these values, $sat(n,k,r) = k(2r-4)n + O(1)$ as $n \to \infty$. Along the way, we disprove a conjecture and answer a question of the first set of authors mentioned above. This work is joint with António Girão and Teeradej Kittipassorn.

2.1 Introduction

Given a graph H, the classical Turán-type extremal problem asks for the maximum number of edges in an H-free graph on n vertices. While the corresponding minimization problem is trivial, it is interesting to determine the minimum number of edges in a maximal H-free graph on n vertices. We say that a graph is *H-saturated* if it is H-free but the addition of an edge joining any two non-adjacent vertices creates a copy of H. The minimum number $sat(n,H)$ of edges in an H-saturated graph on n vertices was first studied in 1949 by Zykov [80] and independently in 1964 by Erdős, Hajnal, and Moon [32] who proved that $sat(n,K_r) = (r-2)(n-1) - \binom{r-2}{2}$. Soon after this, Bollobás [10] determined exactly $sat(n,K_r^{(s)})$ where $K_r^{(s)}$ is the complete s-uniform hypergraph on r vertices. Later, in 1986, Kászonyi and Tuza [54] showed that the saturation number $sat(n,H)$ for a graph H on r vertices is maximized at $H = K_r$, and consequently, $sat(n,H)$ is linear in n for any H. For results on the saturation number, we refer the reader to the survey [33].

This concept can be generalized to the notion of H-saturated subgraphs which are maximal elements of a family of H-free subgraphs of a fixed host graph. A subgraph of a graph G is said to be *H-saturated in G* if it is H-free but the addition of an edge in $E(G)$ joining any two non-adjacent vertices creates a copy of H. The problem of determining the minimum number $sat(G,H)$ of edges in an H-saturated subgraph of G was first proposed in the above mentioned paper of Erdős, Hajnal, and Moon. They conjectured a value for the saturation number $sat(K_{m,n},K_{r,r})$ which was verified independently by Bollobás [11, 12] and Wessel [78, 79]. Very recently, Sullivan and Wenger [75] studied the analogous saturation numbers for tripartite graphs within tripartite graphs and determined $sat(K_{n_1,n_2,n_3},K_{l,l,l})$ for every fixed $l \geq 1$ and every n_1,n_2 and n_3 sufficiently large. Several other host graphs have been considered, including hypercubes [25, 53, 69] and random graphs [58].

In this chapter, we are interested in the saturation number $sat(n,k,r) = sat(K_{k\times n},K_r)$ for $k \geq r \geq 3$ where $K_{k\times n}$ is the complete k-partite graph containing n vertices in each of its k parts. This function was first studied recently by Ferrara, Jacobson, Pfender and Wenger [37] who determined $sat(n,k,3)$ for $n \geq 100$. Later, Roberts [72] showed that $sat(n,4,4) = 18n-21$ for sufficiently large n.

For convenience, we say that a k-partite graph with a fixed k-partition is *K_r-partite-saturated* if it is K_r-free but the addition of an edge joining any two non-adjacent vertices from different parts creates a K_r. Therefore, $sat(n,k,r)$ is the minimum number of edges in a k-partite graph G with n vertices in each part which is K_r-partite-saturated.

Our first result states that $sat(n,k,r)$ is linear in n where the constant $\alpha(k,r)$ in front of n is defined as follows. Given $k \geq r \geq 3$, consider a K_r-partite-saturated k-partite graph G containing an independent set X of size k consisting of exactly one vertex from each part of G. We define $\alpha(k,r)$ to be the minimum number of edges between X and X^c taken over all such G and X.

Theorem 2.1. For $k \geq r \geq 3$,

$$sat(n,k,r) = \alpha(k,r)n + o(n)$$

as $n \to \infty$.

Let us shift our focus to the function $\alpha(k,r)$. The next theorem states what we know about it.

Theorem 2.2. For $k \geq r \geq 3$,

$$(i)\ \ k(2r-4) \leq \alpha(k,r) \leq \begin{cases} (k-1)(4r-k-6) & \text{for } r \leq k \leq 2r-3, \\ (k-1)(2r-3) & \text{for } k \geq 2r-3. \end{cases}$$

$$(ii)\ \ \alpha(k,r) = k(2r-4) \text{ if } \begin{cases} k = 2r-3, \text{ or} \\ k \geq 2r-2 \text{ and } r \equiv 0 \mod 2, \text{ or} \\ k \geq 2r-1 \text{ and } r \equiv 2 \mod 3. \end{cases}$$

(iii) $\alpha(k,3) = 3(k-1)$, $\alpha(4,4) = 18$ and $33 \leq \alpha(5,5) \leq 36$.

(iv) $\alpha(r,r) \geq r(2r-4)+1$ for $r \geq 4$.

The bounds in (i), together with Theorem 2.1, imply that $sat(n,k,r) = O(krn)$, answering a question of Ferrara, Jacobson, Pfender and Wenger [37]. In (ii), we determine exactly $\alpha(k,r)$ for some values of r and every k large enough, allowing us to disprove a conjecture in [37] which states that $sat(n,k,r) = (k-1)(2r-3)n-(2r-3)(r-1)$ for $k \geq 2r-3$ and sufficiently large n. In (iii), we deal with the cases $r = 3,4,5$ which have not been covered by (ii). Finally, (iv) shows that the lower bound in (i), which is attained for certain values of r and k mentioned in (ii), is not tight when $k = r$.

Theorem 2.1 and Theorem 2.2 imply that $sat(n,k,r) = k(2r-4)n + o(n)$ for the values of k and r in (ii). We show that, in this case, the $o(n)$ term can be replaced by a constant.

Theorem 2.3. For $k \geq r \geq 3$,

$$sat(n,k,r) = k(2r-4)n + O(1) \text{ if } \begin{cases} k = 2r-3, \text{ or} \\ k \geq 2r-2 \text{ and } r \equiv 0 \mod 2, \text{ or} \\ k \geq 2r-1 \text{ and } r \equiv 2 \mod 3, \end{cases}$$

as $n \to \infty$.

Now we give a summary of the values of $sat(n,k,r)$ in the case $r = 3,4,5$ which are immediate consequences of the first three results.

Corollary 2.4. (i) $sat(n,k,3) = 3(k-1)n + o(n)$ for $k \geq 3$ and as $n \to \infty$.

$$(ii)\ \ sat(n,k,4) = \begin{cases} 18n + o(n) & \text{for } k = 4, \text{ as } n \to \infty, \\ 4kn + O(1) & \text{for } k \geq 5, \text{ as } n \to \infty. \end{cases}$$

$$(iii)\ \ sat(n,k,5) \begin{cases} \in [33n + o(n), 36n + o(n)] & \text{for } k = 5, \text{ as } n \to \infty, \\ \in [36n + o(n), 40n + o(n)] & \text{for } k = 6, \text{ as } n \to \infty, \\ \in [48n + o(n), 49n + o(n)] & \text{for } k = 8, \text{ as } n \to \infty, \\ = 6kn + O(1) \quad \text{for } k = 7 \text{ or } k \geq 9, \text{ as } n \to \infty. & \square \end{cases}$$

We note that (i) and the first half of (ii) are not the best known results. In fact, Ferrara, Jacobson, Pfender and Wenger [37] proved that $sat(n,k,3) = 3(k-1)n - 6$ for sufficiently large n and Roberts [72] proved that $sat(n,4,4) = 18n - 21$ for sufficiently large n.

Let us give some more definitions which will be used throughout the chapter. For a k-partite $G = V_1 \cup V_2 \cup \cdots \cup V_k$, we refer to each V_i as a *part of* G. We say that an edge (or a non-edge) uv of a k-partite graph is *admissible* if u, v lie in different parts. We say that a non-edge uv of a K_r-free graph is *K_r-saturated* if adding uv to the graph completes a K_r. In other words, a k-partite graph is K_r-partite-saturated if it is K_r-free and every admissible non-edge is K_r-saturated.

The rest of this chapter is organized as follows. Section 2.2 is devoted to the proof of Theorem 2.1. In Section 2.3, we study the function $\alpha(k,r)$ and prove Theorem 2.2(i). In Section 2.4, we prove Theorem 2.2(ii) by describing constructions matching the lower bound $\alpha(k,r) \geq k(2r-4)$ in Theorem 2.2(i). We prove Theorem 2.2(iii), Theorem 2.2(iv) and Theorem 2.3 in Section 2.5, Section 2.6 and Section 2.7 respectively. Finally, we conclude the chapter in Section 2.8 with some open problems.

2.2 Proof of Theorem 2.1

First we show that the upper bound follows easily from the definition of $\alpha(k,r)$.

Proposition 2.5. For every $k \geq r \geq 3$ and any integer $n \geq \alpha(k,r)+1$, we have $sat(n,k,r) \leq \alpha(k,r)n + \alpha(k,r)^2$.

Proof. Let G be a K_r-partite-saturated k-partite graph containing an independent set X of size k consisting of exactly one vertex from each part of G with $e(X, X^c) = \alpha(k,r)$. We may assume that $|X^c| \leq \alpha(k,r)$. Indeed, since there are $\alpha(k,r)$ edges between X and X^c, deleting all the vertices in X^c with no neighbors in X leaves at most $\alpha(k,r)$ vertices in X^c. Note that any admissible non-edge with at least one endpoint in X is still K_r-saturated. We finish by keeping adding admissible edges inside X^c until every admissible non-edge inside X^c is K_r-saturated.

Let $V_1, V_2, \ldots, V_k$ be the parts of G. It follows that $|V_i| = |V_i \cap X| + |V_i \cap X^c| \leq 1 + \alpha(k,r) \leq n$, and so we can modify G to have exactly n

vertices in each part by blowing up the vertex of X in V_i to a class of size $n-|V_i\cap X^c|$ for each i. The resulting graph is K_r-partite-saturated and has exactly n vertices in each of its k parts. Moreover, the number of edges is at most
$\alpha(k,r)n+e(G[X^c])\le\alpha(k,r)n+\alpha(k,r)^2$. □

Now we prove the lower bound $sat(n,k,r)\ge\alpha(k,r)n+o(n)$.

Let $\varepsilon>0$ and let $G=V_1\cup V_2\cup\cdots\cup V_k$ be a K_r-partite-saturated k-partite graph with $|V_i|=n$ for all $i\in[k]$. We shall show that $e(G)\ge\alpha(k,r)n-\varepsilon n$ for all sufficiently large n. Let d be a large natural number to be chosen later. For each i, we partition V_i into $V_i^+=\{v\in V_i : d(x)\ge d\}$ and $V_i^-=\{v\in V_i : d(x)<d\}$. First we show that V_i^+ is small. Since $e(G)\ge\frac{d}{2}|V_i^+|$, we are done unless $|V_i^+|\le\frac{2\alpha(k,r)}{d}n$. Now we show that we can delete a constant number of vertices from $\bigcup_{i=1}^k V_i^-$ to make it independent.

Lemma 2.6. There exists a subset $U\subset\bigcup_{i=1}^k V_i^-$ of size $C_{k,d}$ such that $\left(\bigcup_{i=1}^k V_i^-\right)\setminus U$ forms an independent set in G for some constant $C_{k,d}$.

Let us first show how to finish the proof of Proposition 2.5 using the lemma. For each $1\le i\le k$, let v_i be a vertex of smallest degree in $V_i^-\setminus U$. Since G is a K_r-partite-saturated k-partite graph and $X=\{v_1,v_2,\dots,v_k\}$ is an independent set with exactly one vertex in each part of G, we have $\sum_{i=1}^k d(v_i)\ge\alpha(k,r)$ by the definition of $\alpha(k,r)$. Since $\left(\bigcup_{i=1}^k V_i^-\right)\setminus U$ forms an independent set,

$$\begin{aligned}
e(G)&\ge\sum_{i=1}^k\sum_{v\in V_i^-\setminus U}d(v)\ge\sum_{i=1}^k|V_i^-\setminus U|d(v_i)\ge\sum_{i=1}^k(n-|V_i^+|-|U|)d(v_i)\\
&\ge\left(n-\frac{2\alpha(k,r)}{d}n-C_{k,d}\right)\sum_{i=1}^k d(v_i)\ge\left(n-\frac{2\alpha(k,r)}{d}n-C_{k,d}\right)\alpha(k,r)\\
&=\alpha(k,r)n-\left(\frac{2\alpha(k,r)^2}{d}+\frac{\alpha(k,r)C_{k,d}}{n}\right)n\ge\alpha(k,r)n-\varepsilon n
\end{aligned}$$

by taking d and n sufficiently large. It remains to prove the lemma.

Proof of Lemma 2.6. It is sufficient to show that any matching between V_i^- and V_j^- has size less than 4^{d^2} for all $i \neq j$. Indeed, we can take U to be the endpoints of maximal matchings between V_i^- and V_j^- for all $i \neq j$ and $|U| < 4^{d^2}\binom{k}{2}$.

Suppose for contradiction that $\{x_1y_1, x_2y_2, \dots, x_{4^{d^2}}y_{4^{d^2}}\}$ is a matching of size 4^{d^2} where $X = \{x_1, x_2, \dots, x_{4^{d^2}}\} \subset V_1^-$ and $Y = \{y_1, y_2, \dots, y_{4^{d^2}}\} \subset V_2^-$. The strategy of the proof is to iteratively find vertices $x_{t_1}, x_{t_2}, \dots, x_{t_d}$ of X such that $d(x_{t_i}) \geq i$ for all $1 \leq i \leq d$, which would contradict the fact that $x_{t_d} \in V_1^-$. In fact, we shall find vertices $x_{t_1}, x_{t_2}, \dots, x_{t_d}$ of X such that

(i) there exists a common neighbor of x_{t_i} and y_{t_j} which is not a neighbor of $y_{t_1}, y_{t_2}, \dots, y_{t_{j-1}}$ for all $i > j$.

Clearly, this implies that $d(x_{t_i}) \geq i$ for all $1 \leq i \leq d$. To find such vertices, it is sufficient to find vertices $x_{t_1}, x_{t_2}, \dots, x_{t_d}$ of X satisfying

(ii) x_{t_i} and y_{t_j} are not neighbors for all $i > j$, and

(iii) $N(x_{t_i}) \cap N(y_{t_l}) = N(x_{t_j}) \cap N(y_{t_l})$ for all $i > j > l$.

First we show that (ii) and (iii) imply (i). Let $i > j$. By (ii), $x_{t_i}y_{t_j}$ is a non-edge. Since G is K_r-partite-saturated, there exists a clique W of size $r-2$ in the common neighborhood of x_{t_i} and y_{t_j}. Since $r \geq 3$, we are done by picking a required vertex from W unless each vertex in W is joined to some y_{t_l} with $l < j$. In this case, $W \cup \{x_{t_j}, y_{t_j}\}$ forms a clique of size r, contradicting the fact that G is K_r-free. Indeed, each $w \in W$ belongs to some $N(y_{t_l})$ with $l < j$, and since $w \in N(x_{t_i})$, we must have $w \in N(x_{t_j})$, by (iii).

Now, we find vertices $x_{t_1}, x_{t_2}, \dots, x_{t_d}$ of X satisfying (ii) and (iii). To help us do so, we shall iteratively construct a nested sequence of sets $X \supset X_1 \supset X_2 \supset \cdots \supset X_d$ with $x_{t_i} \in X_i$ for all $2 \leq i \leq d$, satisfying

(iv) x and $y_{t_{i-1}}$ are not neighbors for all $x \in X_i$, and

(v) $N(x) \cap N(y_{t_{i-1}}) = N(x') \cap N(y_{t_{i-1}})$ for all $x, x' \in X_i$.

Clearly, such vertices $x_{t_1}, x_{t_2}, \ldots, x_{t_d}$ satisfy (ii) and (iii). Start with $x_{t_1} = x_1$ and $X_1 = X$. Let $i \leq d$ and suppose that we have found vertices $x_{t_1}, x_{t_2}, \ldots, x_{t_{i-1}}$ and sets $X_1 \supset X_2 \supset \cdots \supset X_{i-1}$ with $x_{t_j} \in X_j$ for all $j < i$, satisfying (iv) and (v). We delete the neighbors of $y_{t_{i-1}}$ from X_{i-1} and partition the remaining vertices into $2^{d(y_{t_{i-1}})} \leq 2^d$ subsets according to their common neighborhood with $y_{t_{i-1}}$. In other words, $X_{i-1} \setminus N(y_{t_{i-1}})$ is partitioned into subsets $\{x : N(x) \cap N(y_{t_{i-1}}) = S\}$ for $S \subset N(y_{t_{i-1}})$. We choose X_i to be such subset of maximum size, i.e. $|X_i| \geq \frac{|X_{i-1}|-d}{2^d}$. Clearly, X_i satisfies (iv) and (v). We then choose x_{t_i} be any vertex in X_i. It remains to prove that $|X_i| > 0$. Recall that $|X_1| = |X| = 4^{d^2}$, and we can see, by induction, that $|X_i| \geq 4^{d(d-i)}$ for $i \leq d$. Indeed,

$$|X_i| \geq \frac{|X_{i-1}|-d}{2^d} \geq \frac{|X_{i-1}|}{4^d} \geq \frac{4^{d(d-i+1)}}{4^d} \geq 4^{d(d-i)}$$

as required. □

2.3 Bounding $\alpha(k,r)$

In this section, we establish a number of results that will help us prove Theorem 2.2. We shall deduce Theorem 2.2(i) at the end of the section.

For $k \geq r \geq 2$ and $1 \leq i \leq k-r+1$, let $\beta_i(k,r)$ be the minimum number of vertices in a K_r-free k-partite graph such that the subgraph induced by any $k-i$ parts contains a K_{r-1}, i.e. the deletion of any i parts does not destroy all the K_{r-1}.

We observe that β_1 and β_2 are useful for bounding α.

Proposition 2.7. For $k \geq r \geq 3$,

$$k\beta_1(k-1,r-1) \leq \alpha(k,r) \leq (k-1)\beta_2(k,r-1).$$

Proof. To prove the lower bound, let G be a K_r-partite-saturated k-partite graph containing an independent set X of size k consisting of exactly one vertex from each part of G. We shall show that $e(X,X^c) \geq k\beta_1(k-1,r-1)$. It is sufficient to show that each vertex in X has degree at least $\beta_1(k-1,r-1)$. Let $x \in X$ and consider the $(k-1)$-partite graph $H = G[N(x)]$. Clearly, it is K_{r-1}-free since G is K_r-free. It remains to show that, for each part U of H, $H \setminus U$ contains a K_{r-2}. If x' is a vertex of X in the corresponding part of U in G then, since the non-edge xx' is K_r-saturated in G, $H \setminus U$ must contain a K_{r-2}. Hence, $|N(x)| = |H| \geq \beta_1(k-1,r-1)$.

For the upper bound, let G_1 be a K_{r-1}-free k-partite graph on $\beta_2(k,r-1)$ vertices such that the subgraph induced by any $k-2$ parts contains a K_{r-2}. Let G_2 be the graph obtained from G_1 by adding one vertex of $X = \{x_1, x_2, \ldots, x_k\}$ to each part of G_1 and joining each x_i to every vertex of G_1 outside its part. By construction, X forms an independent set and $e(X,X^c) = (k-1)\beta_2(k,r-1)$ edges. Note that G_2 is K_r-free since a clique in G_2 contains at most one vertex from X and G_1 is K_{r-1}-free. Now, let G be the graph obtained from G_2 by adding admissible edges inside X^c, until every admissible non-edge inside X^c is K_r-saturated. To conclude that G is K_r-partite-saturated, we need to show that every admissible non-edge inside X is K_r-saturated. Note that, for every pair of distinct vertices $x, x' \in X$, G_1 contains a K_{r-2} not using vertices from the parts containing x and x'. Since x and x' are joined to every vertex outside their parts, the addition of the edge xx' completes a K_r. Hence, $\alpha(k,r) \leq e(X,X^c) = (k-1)\beta_2(k,r-1)$. □

In the next sections, the argument above used in the proof of the lower bound will be used several times. Let us state it as a lemma.

Lemma 2.8. Let G be a k-partite K_r-free graph containing an independent set X of size k consisting of exactly one vertex from each part of G such that the non-edges inside X are K_r-saturated. Then, for each $x \in X$, $G[N(x)]$ is a K_{r-1}-free $(k-1)$-partite graph such that the subgraph induced by any $k-2$ parts contains a K_{r-2}. In particular, $d(x) \geq \beta_1(k-1,r-1)$ for all $x \in X$.

In the next two subsections, we shall bound β_1 from below and β_2 from above.

2.3.1 Upper bounds for β_i

We start with an easy observation which helps us bound β_i from above.

Lemma 2.9. For $k \geq r \geq 3$ and $1 \leq i \leq k-r+1$, $\beta_i(k,r) \leq \beta_i(k-1,r-1)+i+1$.

Proof. Let $H = U_1 \cup U_2 \cup \cdots \cup U_{k-1}$ be a K_{r-1}-free $(k-1)$-partite graph on $\beta_i(k-1,r-1)$ vertices such that the subgraph induced by any $k-i-1$ parts contains a K_{r-2}. We shall construct a K_r-free k-partite graph $G = V_1 \cup V_2 \cup \cdots \cup V_k$ from H with $|G| = |H| + (i+1)$ as follows. First, add new vertices v_1 to U_1, v_2 to U_2, ..., v_i to U_i and v_{i+1} to the new part V_k. This is possible since $k \geq i+2$. Now, join v_{i+1} to every vertex in H and, for every $1 \leq j \leq i$, join v_j to every vertex in $H \setminus U_j$. Clearly, G is K_r-free since H is K_{r-1}-free.

Let $\mathscr{C}$ be a collection of $k-i$ parts of G. It remains to check that the subgraph of G induced by $\mathscr{C}$ contains a K_{r-1}. First, suppose that $V_k \in \mathscr{C}$. By the induction hypothesis, the other $(k-1)-i$ parts $\mathscr{C} \setminus \{V_k\}$ induce a subgraph of H containing a K_{r-2}. Together with $v_{i+1} \in V_k$, they form a K_{r-1} in the subgraph of G induced by $\mathscr{C}$ as required. Now, let us suppose that $V_k \notin \mathscr{C}$. Then $\mathscr{C}$ must contain at least one of $V_1, V_2, \ldots, V_i$. Without loss of generality, we may assume that $\mathscr{C}$ contains V_1. By the induction hypothesis, the other $(k-1)-i$ parts $\mathscr{C} \setminus \{V_1\}$ induce a subgraph of H containing a K_{r-2}. Together with $v_1 \in V_1$, they form a K_{r-1} in the subgraph of G induced by $\mathscr{C}$ as required. □

Lemma 2.9 immediately implies the following upper bound on β_i.

Corollary 2.10. $\beta_i(k,r) \leq (i+1)(r-1)$ for $k \geq r \geq 2$ and $1 \leq i \leq k-r+1$.

Proof. It is clear that $\beta_i(k,2) = i+1$ for $k \geq i+1$ by considering the empty graph on $i+1$ vertices where each vertex is in a different part and the remaining $k-i-1$ parts are empty.

By induction on r and applying Lemma 2.9,

$\beta_i(k,r) \leq \beta_i(k-1,r-1)+i+1 \leq (i+1)(r-2)+i+1 = (i+1)(r-1)$ as required. □

We remark that there is a straightforward construction proving Corollary 2.10 for the case $k \geq (i+1)(r-1)$, namely, a disjoint union of $i+1$ cliques of size $r-1$ where each vertex is in a different part and the remaining $k-(i+1)(r-1)$ parts are empty. Clearly, the deletion of any i parts does not destroy all the K_{r-1}.

Now we prove a better upper bound for $\beta_i(k,r)$ in the case when $i \geq 2$ and $k \geq i(r-1)+1$ by considering the $(r-2)$th power of the cycle $C_{i(r-1)+1}$.

Proposition 2.11. $\beta_i(k,r) \leq i(r-1)+1$ for $k \geq i(r-1)+1$ and $r,i \geq 2$.

Proof. Since $\beta_i(k,r)$ is decreasing in k (by adding empty parts), it is enough to show that $\beta_i(k,r) \leq i(r-1)+1$ for $k = i(r-1)+1$. Let G be the $(r-2)$th power of the cycle $C_{i(r-1)+1}$, i.e. G is a graph on $\mathbb{Z}_{i(r-1)+1}$ where u,v are neighbors if $u-v = 1,2,\ldots,r-2$. We view G as a $(i(r-1)+1)$-partite graph with one vertex in each part. Clearly, G is K_r-free if $i \geq 2$. Note that, after deleting any i vertices of G, there are at least $r-1$ consecutive vertices remaining in $\mathbb{Z}_{i(r-1)+1}$, which form a K_{r-1} as required. □

Proposition 2.11 together with Lemma 2.9 imply a better upper bound than that in Corollary 2.10 for $\beta_2(k,r)$ in the remaining cases, i.e when $k < 2r-1$.

Proposition 2.12. $\beta_2(k,r) \leq 4r-k-2$ for $2 \leq r < k \leq 2r-1$.

Proof. We proceed by induction on $2r-k$. The base case when $2r-k = 1$ follows from Proposition 2.11. Now, suppose that $2r-k \geq 2$. Applying Lemma 2.9,

$$\beta_2(k,r) \leq \beta_2(k-1,r-1)+3 \leq (4(r-1)-(k-1)-2)+3 = 4r-k-2,$$

by the induction hypothesis, since $2r-k > 2(r-1)-(k-1) \geq 1$, □

Let us remark that a similar upper bound for general β_i can be obtained by the same method. We believe that the bound in Proposition 2.12 is, in fact, an equality.

Conjecture 2.13. $\beta_2(k,r) = 4r-k-2$ for $2 \leq r < k < 2r-1$.

For the remaining values of k, we shall see in the next subsection that $\beta_2(k,r) = 2r-1$ for $k \geq 2r-1$.

2.3.2 Determining β_1

We shall show that the upper bound for β_1 given by Corollary 2.10 is an equality. Recall that the clique number of a graph is the order of a maximum clique.

Proposition 2.14. $\beta_1(k,r) = 2(r-1)$ for $k \geq r \geq 2$.

The lower bound is a consequence of the following observation which we will prove below.

Proposition 2.15. Let G be a graph on at most $2s-1$ vertices with clique number s. Then there is a vertex which lies in every K_s of G.

Proof of Proposition 2.14. The upper bound follows from Corollary 2.10. To prove the lower bound, suppose for contradiction that G is a K_r-free k-partite graph on at most $2r-3$ vertices such that the subgraph induced by any $k-1$ parts contains a K_{r-1}. Applying Proposition 2.15 with $s = r-1$, there is a vertex v which lies in every K_{r-1}. In particular, the deletion of the part containing v destroys all the K_{r-1}. Hence, $\beta_1(k,r) \geq 2r-2$. □

Let us remark that Proposition 2.15 is a consequence of the clique collection lemma of Hajnal [50] which states that the sum of the number of vertices in the union and the intersection of a collection of maximum cliques is at least twice the clique number. Our argument below can also be used to give a new proof of Hajnal's clique collection lemma.

Proof of Proposition 2.15. Let $V_1, V_2, \ldots, V_m \subset V(G)$ be the vertex sets of the copies of K_s in G. For a vertex $v \in V(G)$, let $I_v = \{i \in [m] : v \in V_i\}$ be the set of K_s containing v. For a collection $\mathscr{C} \subset \mathscr{P}([m])$ of subsets of $[m]$, let $V_{\mathscr{C}} = \{v \in V(G) : I_v \in \mathscr{C}\}$. Observe that if $\mathscr{C} \subset \mathscr{P}([m])$ is intersecting then $V_{\mathscr{C}}$ induces a clique in G. Indeed, $u, v \in V_{\mathscr{C}}$ are neighbors since $I_u \cap I_v \neq \emptyset$, i.e. there is a clique containing both u and v. Therefore, $|V_{\mathscr{C}}| \leq s$ since G is K_{s+1}-free. The following lemma implies the result.

Lemma 2.16. For $m \geq 3$, there exist intersecting families $\mathscr{C}_1, \mathscr{C}_2, \ldots \mathscr{C}_{m-2} \subset \mathscr{P}([m])$ such that, for $I \subset [m]$, the number of $\mathscr{C}_j$ containing I is $\begin{cases} 0 & \text{if } I = \emptyset \\ |I|-1 & \text{if } I \neq \emptyset, [m] \\ m-2 & \text{if } I = [m]. \end{cases}$

Proof. The proof is by induction on m. For $m = 3$, $\mathscr{C}_1 = \{\{1,2\},\{2,3\},\{3,1\},\{1,2,3\}\}$ satisfies the required property. For $m \geq 4$, suppose by induction that there exist intersecting families $\mathscr{C}_1, \mathscr{C}_2, \ldots, \mathscr{C}_{m-3} \subset \mathscr{P}([m-1])$ satisfying the property. We define $\mathscr{D}_1, \mathscr{D}_2, \ldots, \mathscr{D}_{m-2} \subset \mathscr{P}([m])$ as follows. For $1 \leq j \leq m-3$, let

$$\mathscr{D}_j = \mathscr{C}_j \cup \{I \cup \{m\} : I \in \mathscr{C}_j\}$$

and

$$\mathscr{D}_{m-2} = \{I \subset [m] : m \in I \text{ and } |I| \geq 2\} \cup \{[m-1]\}.$$

It is easy to check that $\mathscr{D}_1, \mathscr{D}_2, \ldots, \mathscr{D}_{m-2}$ satisfy the required property. $\square$

Let us deduce the result. This is trivial when $m = 1,2$ so we may assume that $m \geq 3$. Observe that

$$\sum_{i=1}^{m} |V_i| = \left|\bigcup_{i=1}^{m} V_i\right| + \sum_{j=1}^{m-2} |V_{\mathscr{C}_j}| + \left|\bigcap_{i=1}^{m} V_i\right|.$$

Indeed, a vertex v is counted on both sides $|I_v|$ times by the lemma. Using $|V_i| = s$, $|\bigcup_{i=1}^{m} V_i| \leq 2s-1$ and $|V_{\mathscr{C}_j}| \leq s$, we have

$$ms \leq (2s-1) + (m-2)s + \left|\bigcap_{i=1}^{m} V_i\right|$$

i.e. $|\bigcap_{i=1}^{m} V_i| \geq 1$ as required. $\square$

We remark that the fact that $\beta_1(k,r) = 2(r-1)$ allows us to show that the upper bound for $\beta_2(k,r)$ when $k \geq 2r-1$ in Proposition 2.11 is an equality.

Corollary 2.17. $\beta_2(k,r) = 2r-1$ for $k \geq 2r-1$ and $r \geq 2$.

Proof. Observe that $\beta_i(k,r) \geq \beta_{i-1}(k-1,r)+1$. Indeed, if G is a K_r-free k-partite graph on $\beta_i(k,r)$ vertices such that the subgraph induced by any $k-i$ parts contains a K_{r-1}, then, by deleting a non-empty part of G, we obtain a K_r-free $(k-1)$-partite graph such that the subgraph induced by any $(k-1)-(i-1)$ parts contains a K_{r-1}. This graph must contains at least $\beta_{i-1}(k-1,r)$ vertices and therefore, $|G|-1 \geq \beta_{i-1}(k-1,r)$.

Hence, $\beta_2(k,r) \geq \beta_1(k-1,r)+1 = 2(r-1)+1 = 2r-1$ by Proposition 2.14. □

2.3.3 Proof of Theorem 2.2(i)

The lower bound follows from Proposition 2.7 and Proposition 2.14. The upper bound follows from Proposition 2.7, Proposition 2.12 and Corollary 2.17. □

2.4 Proof of Theorem 2.2(ii)

For $k = 2r-3$, we are done since the lower and upper bounds in Theorem 2.2(i) match, i.e. $\alpha(k,r) = k(2r-4) = (k-1)(2r-3)$.

Now we shall describe constructions that match the lower bound $\alpha(k,r) \geq k(2r-4)$ in Theorem 2.2(i) for the cases when ($k \geq 2r-2$ and r is even) and ($k \geq 2r-1$ and $r = 2$ mod 3), i.e. a K_r-partite-saturated k-partite graph G containing an independent set X of size k consisting of exactly one vertex from each part of G with $e(X,X^c) = k(2r-4)$. Lemma 2.8 tells us that such graph must satisfy $d(x) = 2r-4$, for all $x \in X$.

Note that we do not have to worry about making the admissible non-edges inside X^c, K_r-saturated since we can keep adding admissible edges inside X^c until every admissible non-edge inside X^c is K_r-saturated.

Let $p \in \{2,3\}$ be a divisor of $r-2$. First we shall construct such k-partite graph G, for $k = 2r-4+p$. We define $X = \{x_1,x_2,\ldots,x_k\}$ and $X^c = \{y_1,y_2,\ldots,y_k\}$, where the parts of G are $\{x_i,y_i\}$, for $i = 1,2,\ldots,k$. There are no edges inside X. Let y_iy_j be an edge iff i,j are not consecutive elements of the circle $\mathbb{Z}_k$, and so $G[X^c]$ is the graph K_k minus a cycle C_k. Let x_iy_j is an edge iff $i \neq j \mod \frac{k}{p}$, i.e. x_i is joined to all but p equally spaced y_j. We claim that G satisfies the required properties.

Clearly, we have $d(x) = k - p = 2r - 4$ for all $x \in X$ and $e(X, X^c) = k(2r-4)$. Let us verify that G is K_r-free. A clique inside X^c is a set of non-consecutive elements of $\mathbb{Z}_k$, and so a largest clique inside X^c has size $\lfloor \frac{k}{2} \rfloor = r-1$ for $p \in \{2,3\}$. Since a clique which is not inside X^c can contain at most one vertex of X, it remains to check that the neighborhood of each x_i does not contain a clique of size $r-1$. Viewing X^c as a circle, $N(x_i)$ consists of p segments of the circle, each of size $\frac{2r-4}{p}$, separated by gaps of size one. Since $\frac{2r-4}{p}$ is even, a largest clique in $N(x_i)$ has size $\frac{p(2r-4)}{2p} = r-2$.

It remains to show that the admissible non-edges inside X, and those between X and X^c are K_r-saturated. Let x_iy_j be an admissible non-edge, and so $j = i \pm \frac{k}{p}$ in $\mathbb{Z}_k$. Clearly, $N(x_i)$ contains $r-2$ vertices which form a non-consecutive set of the circle with y_j. Therefore, there exists a K_{r-2} in the common neighborhood of x_i and y_j as required. Now let x_ix_j be an admissible non-edge. Then the common neighborhood of x_i and x_j consists of $2p$ segments of the circle separated by gaps of size one such that they form p pairs where the sum of the sizes of each pair is $\frac{2r-4}{p} - 1$, and so each pair consists of a segment of even size and a segment of odd size. Therefore, a largest non-consecutive set in $N(x_i) \cap N(x_j)$ has size $\frac{p(2r-4)}{2p} = r-2$. Hence, there exists a K_{r-2} in $N(x_i) \cap N(x_j)$ as required.

We have constructed such k-partite graph G_k for $k = 2r-4+p$. Let us obtain G_k for $k > 2r-4+p$ from G_{2r-4+p} by blowing up x_1 to a class $\{x_1\} \cup \{x_i : 2r-3+p \leq i \leq k\}$ of size $k-(2r-4+p)+1$ where each copy of x_1 (not including itself) forms a part of G_k of size one. Clearly, we have $d(x) = 2r-4$ for all $x \in X = \{x_1, x_2, \ldots, x_k\}$ and $e(X, X^c) = k(2r-4)$. Since G_{2r-4+p} is K_r-free, so is G_k.

It remains to check that the admissible non-edges inside X, and those between X and X^c are K_r-saturated. Any admissible non-edge inside X which is not inside the blow up class of x_1 is K_r-saturated by the same property of G_{2r-4+p}. Any admissible non-edge inside the blow up class of x_1 is K_r-saturated since $N(x_1)$ contains a K_{r-2} by the construction of G_{2r-4+p}. Any admissible non-edge x_iy_j where $j \neq 1$ or ($j = 1$ and $i \leq 2r-4+p$), is K_r-saturated by the same property of G_{2r-4+p}. Any admissible non-edge x_iy_j where $j = 1$ and $2r-3+p \leq i \leq k$, is K_r-saturated since $N(x_1) \cap N(y_1)$ contains a K_{r-2} by the construction of G_{2r-4+p}. □

2.5 Proof of Theorem 2.2(iii)

In this section, we study $\alpha(k,r)$ for $r = 3,4,5$. The values of $\alpha(k,3)$ and $\alpha(k,4)$ are completely determined while the values of $\alpha(k,5)$ are unknown for $k = 5,6,8$.

2.5.1 The function $\alpha(k,3)$

We shall prove that $\alpha(k,3) = 3(k-1)$ for $k \geq 3$. The upper bound follows from Theorem 2.2(i). Let us prove the lower bound.

Let $G = V_1 \cup V_2 \cup \cdots \cup V_k$ be a K_3-partite-saturated k-partite graph G containing an independent set $X = \{x_1, x_2, \ldots, x_k\}$ with $x_i \in V_i$ for all i. By Lemma 2.8, for all i, the deletion of any part of G does not destroy all vertices of $N(x_i)$, i.e. x_i is joined to at least two parts of G. Suppose for contradiction that $e(X, X^c) < 3(k-1)$, i.e. X contains at least four vertices of degree 2, say x_1, x_2, x_3, x_4. Let $y_i \in V_i$ and $y_j \in V_j$ with $1 < i < j \leq k$ be the neighbors of x_1, and so y_i and y_j are not neighbors otherwise $x_1y_iy_j$ forms a triangle. Since $\{2,3,4\} \setminus \{i,j\} \neq \emptyset$, we may assume that $i, j \neq 2$, i.e. x_1, x_2, y_i, y_j are from different parts of G. Since any pair in X forms a K_3-saturated non-edge in G, they have a common neighbor. So x_1 and x_2 have a common neighbor, say y_i.

First we suppose that x_2y_j is a non-edge. Then x_2 and y_j have a common neighbor $y_l \in V_l$. Since y_i and y_j are not neighbors, $l \neq i$. We observe that x_iy_j are neighbors since x_1 and x_i have a common neighbor and $N(x_1) = \{y_i, y_j\}$. Similarly, x_iy_l are neighbors since x_2 and x_i have a common neighbor and $N(x_2) = \{y_i, y_l\}$. We obtain a contradiction by observing that $x_iy_jy_l$ forms a triangle.

Now, suppose that x_2y_j is an edge, and so $N(x_1) = N(x_2) = \{y_i, y_j\}$. Then x_iy_j are neighbors since x_1 and x_i have a common neighbor. Similarly, x_jy_i are neighbors. We know that x_i and x_j have a common neighbor y_l with $l \neq i, j$. Then either $l \neq 1$ or $l \neq 2$, say $l \neq 1$. Since the non-edge x_1y_l is K_3-saturated, y_l is joined to either y_i or y_j. This implies a contradiction that either $x_jy_iy_l$ or $x_iy_jy_l$ forms a triangle. □

2.5.2 The function $\alpha(k,4)$

As a consequence of Theorem 2.2(ii), we obtain that $\alpha(k,4) = 4k$ for $k \geq 5$. For the remaining case $k = 4$, we have the bounds $16 \leq \alpha(4,4) \leq 18$ from Theorem 2.2(i). We shall show that $\alpha(4,4) = 18$.

Consider the family of graphs appearing in the definition of $\alpha(r,r)$. Let $G = V_1 \cup V_2 \cup \cdots \cup V_r$ be an K_r-partite-saturated r-partite graph G containing an independent set $X = \{x_1, x_2, \ldots, x_r\}$ with $x_i \in V_i$ for all i. We shall establish some properties of G which will be useful in this subsection, the next subsection and Section 2.6.

We say that a vertex $y \in X^c$ is *i-special* if y is the only neighbor of x_i in the part of G containing y. The *special degree* of a vertex $y \in X^c$ is the number of $i \in [r]$ such that y is i-special. We say that a vertex $y \in X^c$ is *special* if the special degree of y is at least one. Let us make some easy observations regarding the special vertices.

Lemma 2.18. Let $G = V_1 \cup V_2 \cup \cdots \cup V_r$ be an K_r-partite-saturated r-partite graph G containing an independent set $X = \{x_1, x_2, \ldots, x_r\}$ with $x_i \in V_i$ for all i. The following hold for $r \geq 4$.

(i) A special vertex $y_i \in V_i$ is joined to every vertex of X except x_i.

(ii) Each V_i contains at most one special vertex.

(iii) If $y_i \in V_i$ is i'-special and $y_j \in V_j$ is j'-special with $i' \neq j$ and $j' \neq i$ then y_iy_j is an edge.

(iv) The number of vertices of special degree at least 2 is at most $r-2$.

(v) If $y_i \in V_i$ is i'-special and $y_j \in V_j$ with $j \neq i, i'$ then y_j is joined to either y_i or $x_{i'}$.

(vi) For a special vertex $y_i \in V_i$, there exist parts V_j and V_l where i, j, l are distinct such that $N(x_i) \cap V_j$ and $N(x_i) \cap V_l$ both contain a non-neighbor of y_i.

Proof. (i) Let $y_i \in V_i$ be i'-special and let $j \neq i, i'$. Since the non-edge $x_{i'}x_j$ is K_r-saturated, the common neighborhood of $x_{i'}$ and x_j contains a K_{r-2} consisting of one vertex from each part of $G \setminus \left(V_{i'} \cup V_j\right)$. Then y_i is in this K_{r-2} since y_i is the only neighbor of $x_{i'}$ in V_i, and so y_i is joined to x_j.

(ii) Suppose for contradiction that V_i contains two special vertices y_i and z_i where y_i is i'-special. Then, by (i), $x_{i'}$ is joined to both y_i and z_i contradicting the fact that y_i is the only neighbor of $x_{i'}$ in V_i.

(iii) First, suppose that $i' \neq j'$. Since the non-edge $x_{i'}x_{j'}$ is K_r-saturated, the common neighborhood of $x_{i'}$ and $x_{j'}$ contains a K_{r-2} consisting of one vertex from each part $G \setminus \left(V_{i'} \cup V_{j'}\right)$. Since y_i is the only neighbor of $x_{i'}$ in V_i and y_j is the only neighbor of $x_{j'}$ in V_j, both y_i and y_j lie in this K_{r-2}. In particular, y_iy_j is an edge.

Now, suppose that $i' = j'$. We can pick $l \neq i, j, i'$ because $r \geq 4$. Since the non-edge $x_{i'}x_l$ is K_r-saturated, the common neighborhood of $x_{i'}$ and x_l contains a K_{r-2} consisting of one vertex from each part of $G \setminus (V_{i'} \cup V_l)$. Since y_i is the only neighbor of $x_{i'}$ in V_i and y_j is the only neighbor of $x_{i'}$ in V_j, both y_i and y_j lie in this K_{r-2}. In particular, y_iy_j is an edge.

(iv) Suppose for contradiction that there exist vertices $y_1, y_2, \ldots, y_{r-1}$ of special degree at least 2. By (ii), they lie in different parts of G, say $y_i \in V_i$ for $1 \leq i \leq r-1$. We claim that they form a K_{r-1} which would be a contradiction since, together with x_r, they form a K_r by (i). Now we show that any y_iy_j is an edge. Since y_i and y_j have special degree at least 2, there exist $i' \neq j$ and $j' \neq i$ such that y_i is i'-special and y_j is j'-special. Therefore, y_iy_j is an edge by (iii).

(v) Suppose that $x_{i'}y_j$ is a non-edge. Then the common neighborhood of $x_{i'}$ and y_j

contains a K_{r-2} consisting of one vertex from each part of $G \setminus (V_{i'} \cup V_j)$. Then y_i is in this K_{r-2} since y_i is the only neighbor of $x_{i'}$ in V_i, and so y_i is joined to y_j.

(vi) Suppose for contradiction that there exists $j \in [r] \setminus \{i\}$ such that $y_i \in V_i$ is joined to every vertex in $N(x_i) \cap V_l$ for all $l \neq i, j$. Since the non-edge $x_i x_j$ is K_r-saturated, the common neighborhood of x_i and x_j contains a K_{r-2} consisting of one vertex from each part of $(G \setminus X) \setminus (V_i \cup V_j)$. We obtain a contradiction by observing that this K_{r-2}, together with x_j and y_i, form a K_r. Indeed, by assumption, this K_{r-2} is also in the neighborhood of y_i and $x_j y_i$ is an edge by (i). □

Now we are ready to show that $\alpha(4,4) \geq 18$. Suppose for contradiction that $\alpha(4,4) \leq 17$, i.e. there exists a K_4-partite-saturated 4-partite graph $G = V_1 \cup V_2 \cup V_3 \cup V_4$ containing an independent set $X = \{x_1, x_2, x_3, x_4\}$ with $x_i \in V_i$ for all i such that $\sum_{i=1}^{4} d(x_i) \leq 17$. By Lemma 2.8, $d(x_i) \geq \beta_1(3,3) = 4$ and each x_i has some neighbor in V_j for $j \neq i$. Therefore, there are at least three vertices of degree 4 and possibly one of degree 5. Since a vertex of degree 4 in X creates at least two special vertices and a vertex of degree 5 in X creates at least one special vertex, the sum of the special degrees of the vertices in X^c is at least $2+2+2+1 = 7$. By Lemma 2.18(iv), there is a vertex of special degree 3, say $y_1 \in V_1$.

For $i = 2,3,4$, since y_1 is i-special, x_i has at least three neighbors in $N(y_1) \cup \{y_1\}$, each in a different part of G, by Lemma 2.8. On the other hand, y_1 has at least two non-neighbors, say $y_2 \in V_2$ and $y_3 \in V_3$, by Lemma 2.18(vi). By Lemma 2.18(v), $x_i y_2$ is an edge for $i \neq 2$ and $x_i y_3$ is an edge for $i \neq 3$. So x_4 has five neighbors, i.e. y_2, y_3 and three vertices in $N(y_1) \cup \{y_1\}$, and $d(x_1) = d(x_2) = d(x_3) = 4$. Since x_2 has four neighbors including y_3 and it has some neighbor in $(N(y_1) \cup \{y_1\}) \cap V_j$ for each $j = 1,3,4$, it has exactly one neighbor in V_4, say y_4. Similarly, x_3 has exactly one neighbor in V_4 which has to be the same vertex y_4 by Lemma 2.18(ii).

We obtain a contradiction by observing that $x_1 y_2 y_3 y_4$ forms a K_4. First, note that $x_1 y_4$ is an edge by Lemma 2.18(i). Now y_4 is not 1-special otherwise y_4 would have special degree 3 and by repeating the argument above with y_1 replaced by y_4, we could deduce that x_1, x_2, or x_3 had degree 5. Therefore, the neighbors of x_1 are y_2, y_3, y_4 and a vertex in V_4. Since y_2, y_3 are both 1-special and y_4 is $2,3$-special, $y_2y_3y_4$ forms a triangle by Lemma 2.18(iii). □

2.5.3 The function $\alpha(k,5)$

As a consequence of Theorem 2.2(i) and (ii), we obtain that

$$\alpha(k,5) = 6k \quad \text{for } k = 7 \text{ or } k \geq 9,$$
$$30 \leq \alpha(5,5) \leq 36,$$
$$36 \leq \alpha(6,5) \leq 40,$$
$$48 \leq \alpha(8,5) \leq 49.$$

We shall improve the lower bound for $\alpha(5,5)$ to 33.

Suppose for contradiction that $\alpha(5,5) \leq 32$, i.e. there exists a K_5-partite-saturated 5-partite graph $G = V_1 \cup V_2 \cup V_3 \cup V_4 \cup V_5$ containing an independent set $X = \{x_1, x_2, x_3, x_4, x_5\}$ with $x_i \in V_i$ for all i such that $\sum_{i=1}^{5} d(x_i) \leq 32$. Write Y_i for $V_i \setminus \{x_i\}$. By Lemma 2.8, $d(x_i) \geq \beta_1(4,4) = 6$ and each x_i has some neighbor in V_j for $j \neq i$. Therefore, there are either four vertices in X of degree 6 or there are three vertices of degree 6 and two of degree 7. Since a vertex of degree 6 in X creates at least two special vertices and a vertex of degree 7 in X creates at least one special vertex, the sum of the special degrees of the vertices in X^c is at least 8, and hence, there exists a vertex of special degree at least two. Let i be such that there is a special vertex $y \in Y_i$ with special degree $d_s(y)$ at least two where $(d(x_i), d_s(y))$ is maximum in lexicographical order[1]. Without loss of generality we can assume that $i = 1$. Let $N = N(y) \setminus X$. By Lemma 2.18(vi), x_1 has two neighbors, say y_2, y_3, belonging to two distinct parts of G, different from V_1, which are

[1] We say that $(a,b) \preccurlyeq (c,d)$ if $a < c$ or $a = c$ and $b \leq d$, where $\preccurlyeq$ denotes the lexicographical order relation.

non-neighbors of y. Without loss of generality, we can assume that $y_2 \in Y_2$ and $y_3 \in Y_3$.

For a pair of non-adjacent vertices $u, v \in G$ and $S \subset G$, we say that S is *uv-saturating* if adding the edge of uv to G creates a copy K of K_5 such that $S \subseteq K$. If $S = \{z\}$ then we simply say that z is uv-saturating. Notice that if S is uv-saturating then S induces a clique.

In the rest of the proof, we shall repeatedly use the following lemma.

Lemma 2.19. Given $i \in \{2,3,4,5\}$ the following hold.

(i) If $j \in \{2,3,4,5\} \setminus \{i\}$ then x_i has a neighbor in $V_j \cap N$. In particular, $d_N(x_i) \geq 3$.

(ii) If y is i-special then x_i is adjacent to y_j for every $j \in \{2,3\} \setminus \{i\}$.

(iii) If y is $x_i x_j$-saturating, for every $j \in \{2,3,4,5\} \setminus \{i\}$, then $d_N(x_i) \geq 4$.

(iv) If y is i-special or $d_s(y) \geq 3$ then $d_N(x_i) \geq 4$.

(v) If y is $2,3$-special and $i \in \{4,5\}$, then $d(x_i) \geq 7$.

(vi) If $i \in \{2,3\}$ and there are p vertices in $X \setminus \{x_1\}$ all of which have neighbors in $Y_i \setminus N$ then there is no vertex in V_i with special degree bigger than $\max\{1, 3-p\}$.

(vii) $Y_3 \cup Y_4 \subset N$.

Proof. (i) Observe that we can choose $k \in \{2,3,4,5\} \setminus \{i, j\}$ such that y is either i-special or k-special. Since there must be a triangle in the common neighborhood of x_i and x_j which uses y, we have that the remaining two vertices belong to N. Hence x_i has a neighbor in $N \cap x_j$.

(ii) This follows directly from Lemma 2.18(v).

(iii) We shall show that $d_N(x_i) \geq \beta_1(3,3) = 4$. Take any $j \in \{2,3,4,5\} \setminus \{i\}$. Since y is $x_i x_j$-saturating then there is an edge in the common neighborhood of x_i and x_j in $N \setminus (V_i \cup V_j)$. Observe that the common neighborhood of x_i and y cannot contain a K_3, hence $d_N(x_i) \geq \beta_1(3,3) = 4$.

(iv) Take any $j \in \{2,3,4,5\} \setminus \{i\}$. Since y is either i- or j-special, it follows that y is $x_i x_j$-saturating. Hence, by (ii), $d_N(x_i) \geq 4$.

(v) Without loss of generality we can assume that $i = 4$. If y is also 4-special then it follows from (ii) and (iv) that $d_N(x_4) \geq 4$ and x_4 is adjacent to y, y_2, y_3, therefore $d(x_4) \geq 7$. Hence we can assume that y is not 4-special. Suppose for contradiction that $d(x_4) = 6$. From (i), we have that $d_N(x_4) \geq 3$ and since y is not 4-special we have that $d_{Y_1}(x_4) \geq 2$. Moreover, x_4 has to have at least one neighbor not in $Y_1 \cup N$ as otherwise there would be a copy of K_5 in G, as seen by considering the non-edge x_1x_4. Therefore, $d(x_4) = d_{Y_1}(x_4) + d_N(x_4) + |N(x_4) \setminus (Y_1 \cup N)| \geq 3+2+1 = 6 = d(x_4)$. Hence, $d_{Y_1}(x_4) = 3$, $d_N(x_4) = 4$ and $|N(x_4) \setminus (Y_1 \cup N)| = 1$. We shall obtain a contradiction by finding a copy of K_5 in the graph G.

Suppose $\{z_1, z_2, z_3\}$ is x_4x_5 saturating, with $z_i \in V_i$. We claim that $y \neq z_1$ and $\{z_2, z_3\} \not\subseteq N$. Suppose for contradiction that it is not the case. If y is x_4x_5-saturating then from (iii) we have that $d_N(x_4) \geq 4$ hence we obtain a contradiction. We can therefore assume that y is not x_4x_5-saturating and hence $z_1 \neq y$. Whence $z_2, z_3 \in N$. Recall that $\{z_1, z_2, z_3\}$ form a triangle and therefore there is an edge between z_2, z_3. By assumption z_2 and z_3 are neighbors of y, hence y, z_2, z_3 form a triangle, and therefore y is x_4x_5-saturating since y, z_2, z_3 belong to the common neighborhood of x_4 and x_5, which contradicts the assumption that y is not x_4x_5-saturating.

Without loss of generality we can assume that $z_2 \notin N$. Using (i), we can therefore suppose that $N(x_4) \cap Y_1 = \{y, z_1\}$, $N(x_4) \cap Y_2 = \{w, z_2\}$, $N(x_4) \cap Y_3 = \{z_3\}$ and $N(x_4) \cap Y_5 = \{z_5\}$, for some $w, z_3, z_5 \in N$. We shall obtain a contradiction by observing that z_1, z_2, z_3, x_4, z_5 form a copy of K_5. First we claim that $\{z_2, z_3, z_5\}$ is x_1x_4-saturating. Indeed, there must be a triangle in the common neighborhood of x_1 and x_4, with one vertex in each V_3, V_4, V_5. There are only two candidates for the triangle: z_2, z_3, z_5 or w, z_3, z_5. It cannot be w, z_3, z_5 since they are all neighbors of y, hence y, w, z_3, x_4, z_5 would form a copy of K_5. Hence we must have that the set $\{z_2, z_3, z_5\}$ is x_1x_4-saturating. Now, since x_4 is not adjacent to y_3, and y_3 is not adjacent to y we must have an edge between z_1 and z_5. Indeed, there must be a triangle in the common neighborhood of x_4 and y_3 with a vertex in each V_1, V_2, V_3. Since x_4 has only one neighbor in V_5, i.e. z_5, and x_4 and y_3 have only one common neighbor in V_1, i.e. z_1, we must have an edge between z_1 and z_5.

Therefore we have that z_1, z_2, z_3 form a triangle, z_2, z_3, z_5 form a triangle, and z_1, z_5 are adjacent. It easy to see now that z_1, z_2, z_3, x_4, z_5 form a copy of K_5.

(vi) Let v be a special vertex in $V_2 \cup V_3$, say in V_2. First observe that if v is 1-special then x_3, x_4, x_5 are all adjacent to $y_2 \in Y_2 \setminus N$. On the other hand, it follows from (i) that x_3, x_4, x_5 all have neighbors in $N \cap Y_2$ hence they all have degree at least 2 in Y_2. It follows that v has special degree 1. If we assume that v is not 1-special then v has special degree at most $3 - p$, since p of the vertices x_3, x_4, x_5 have degree 2 in Y_2.

(vii) Assume for contradiction that there is v, say in $Y_4 \setminus N$. Observe that if y is i-special then it follows from (ii) and (iv) that $d(x_i) \geq 7$, hence if $d_s(y) \geq 3$ we obtain contradiction by finding three vertices in X of degree at least 7. Therefore we can assume that $d_s(y) = 2$.

If y is $5, i$-special, then from (ii) and (iv) we have that $d(x_5) \geq 8$ and $d(x_i) \geq 7$ hence again we obtain a contradiction. Therefore we can assume that y is not 5-special. If y is $2,3$-special then $d(x_2), d(x_3) \geq 7$ and from (iv) we have that $d(x_4), d(x_5) \geq 7$. Hence we can assume that y is $2,4$-special or $3,4$-special. Suppose that the former is the case. Then $d(x_2), d(x_4) \geq 7$. It follows that $d(x_1) = 6$. Therefore by maximality (x_1, y) and from (v) we have that every vertex in $Y_2 \cup Y_3 \cup Y_4$ has special degree at most 1 and no vertex in Y_5 has special degree bigger than 2. Which gives a contradiction since the sum of special degree is then at most 7. □

We are now ready to finish showing that $\alpha(5,5) \geq 33$. We consider several cases depending on the special degree of y.

Case 1. $d_s(y) = 4$

Consider the 4-partite graph $H = G[N(y)]$ with an independent set $X' = \{x_2, x_3, x_4, x_5\}$. Clearly, H is K_4-free since G is K_5-free. We modify H by keeping adding admissible edges inside $H \setminus X'$ until every admissible non-edge inside $H \setminus X'$ is K_4-saturated. We claim that H is K_4-partite-saturated, which would imply that $e(X', H \setminus X') \geq \alpha(4,4) = 18$ by the previous subsection. It remains to show that the admissible non-edges with at least one endpoint in X' are K_4-saturated.

Consider the non-edge $x_i y_j$ with $y_j \in V_j \cap H$ (possibly $y_j = x_j$) and distinct $2 \leq i, j \leq 5$. Since the non-edge $x_i y_j$ is K_5-saturated in G, the common neighborhood in G of x_i and y_j contains a K_3 consisting of one vertex from each part of $G \setminus (V_i \cup V_j)$. Since y is i-special, this K_3 must contain y, and so the common neighborhood in H of x_i and y_j contains a K_2, i.e. $x_i y_j$ is K_4-saturated in H as required.

Recall that y has two non-neighbors, $y_2 \in V_2$ and $y_3 \in V_3$. By Lemma 2.18(v), $x_i y_2$ is an edge for $i \neq 2$ and $x_i y_3$ is an edge for $i \neq 3$. We shall partition the edges between X and X^c as follows:

$$\begin{aligned} e(X, X^c) &\geq e(X', H \setminus X') + d(x_1) + e(X, y) + e(X', y_2) + e(X', y_3) \\ &\geq 18 + 6 + 4 + 3 + 3 = 34, \end{aligned}$$

contradicting the assumption.

Case 2. $d_s(y) = 3$

If y is $4,5$-special then from Lemma 2.19(ii) and 2.19(iii) we have that $d(x_4), d(x_5) \geq 7$. Otherwise y is $2,3$-special and hence it follows from Lemma 2.19(v) that $d(x_4), d(x_5) \geq 7$. We shall obtain a contradiction by showing that $d(x_1) \geq 7$, hence showing that there are three vertices in X with degrees at least 7, which is against an assumption made in the beginning of the subsection. It follows from Lemma 2.19(vi) with $p \geq 2$, that the sum of special degrees in $Y_2 \cup Y_3$ is at most 2. Since the sum of special degrees is at least 8, it follows that there is a special vertex in $Y_4 \cup Y_5$ with special degree at least 2. Therefore from the maximality of $d(x_1)$ we have that $d(x_1) \geq 7$.

Case 3. $d_s(y) = 2$

We split this case into three subcases.

Case 3.1. y is $2,3$-special

It follows from Lemma 2.19(v) that $d(x_4), d(x_5) \geq 7$. We shall obtain a contradiction by showing that $d(x_1) \geq 7$, hence showing that there are three vertices in X with degrees at least 7, which is against an assumption made in the beginning of the subsection. It follows from Lemma 2.19(vi) that the sum of special degrees in $Y_2 \cup Y_3$ is at most 2. Since the sum of special degrees is at least 8, it follows that there is a special vertex in $Y_4 \cup Y_5$ with special degree at least 2. Therefore from the maximality of $d(x_1)$ we have that $d(x_1) \geq 7$.

Case 3.2. y is $4,5$-special

It follows from Lemma 2.19(ii) and 2.19(iv) that $d(x_4), d(x_5) \geq 7$. We shall obtain a contradiction by showing that $d(x_1) \geq 7$, hence showing that there are three vertices in X with degrees at least 7, which is against an assumption made in the beginning of the subsection. It follows from Lemma 2.19(vi) that the sum of special degrees in $Y_2 \cup Y_3$ is at most 3. Since the sum of special degrees is at least 8, it follows that there is a special vertex in $Y_4 \cup Y_5$ with special degree at least 2. Therefore from the maximality of $d(x_1)$ we have that $d(x_1) \geq 7$.

Case 3.3. y is neither $2,3$-special nor $4,5$-special

Without loss of generality we can assume that y is $2,4$-special. It follows from Lemma 2.19(ii) and 2.19(iv) that $d(x_4) \geq 7$ and from Lemma 2.19(vi) with $p \geq 2$ that there is no special vertex in $Y_2 \cup Y_3$ with special degree bigger than 1. Hence there is either a vertex in Y_4 with special degree at least 2 or a vertex in Y_5 with special degree at least 3. Therefore we can assume that $d(x_1) = 7$ as otherwise we obtain a contradiction to the maximality of $(d(x_1), d_s(y))$.

We shall obtain a contradiction by showing that at least one of x_2, x_3 or x_5 has degree at least 7, thus finding three vertices with degree at least 7. Suppose $d(x_2) = d(x_3) = d(x_5) = 6$. Observe that if there is a vertex in X_4 of special degree bigger than 2 then we obtain a contradiction to the maximality of $(d(x_1), d_s(y))$. Therefore there are two vertices in X with at least two neighbors in X_4. Suppose that $i \in \{3,5\}$ and x_i has at least two neighbors in X_4. Then it follows from Lemma 2.19(i) that $d_N(x_i) \geq 4$, and hence x_i has degree at least 7 as x_i has at least three neighbors outside N. We can therefore assume that x_3 and x_5 have only one neighbor in X_4. For the same reason we can assume that x_3 has only one neighbor in Y_5. If x_2 has two neighbors in Y_5 then $d_N(x_2) \geq 5$ and therefore $d(x_2) \geq 7$. Hence we can assume that there is $z_5 \in Y_5$ which is $2,3$-special.

Suppose $\{z_1, z_2, z_4\}$ is x_3x_5-saturating, with $z_i \in V_i$. We claim that $y \neq z_1$ and $z_2 \notin N$. Suppose for contradiction that it is not the case. If y is x_3x_5-saturating then from (iii) we have that $d_N(x_3) \geq 4$ hence we obtain a contradiction. We can therefore assume that y is not x_3x_5-saturating and hence $z_1 \neq y$. Whence $z_2 \in N$. Observe that by Lemma 2.19(vii) we have $z_4 \in N$. Recall that $\{z_1, z_2, z_4\}$ form a triangle and therefore there is an edge between z_2, z_4. By assumption z_2 and z_4 are neighbors of y, hence y, z_2, z_4 form a triangle, and therefore y is x_3x_5-saturating since y, z_2, z_4 belong to the common neighborhood of x_3 and x_5, which contradicts the assumption that y is not x_3x_5-saturating.

We shall obtain a contradiction by showing that z_1, z_2, x_3, z_4, z_5 form a copy of K_5. Indeed, by assumption $\{z_1, z_2, z_4\}$ is x_3x_5-saturating and similar analysis to the one made in the proof of Lemma 2.19(v) shows that $\{z_2, z_4, z_5\}$ is x_1x_3-saturating. Since y is 2-special it follows that x_2 is not adjacent to z_1, and moreover z_5, as the only neighbor of x_2 in Y_5, is x_2z_1-saturating, and therefore there is an edge between x_2 and z_5. Hence we have that z_2, z_4, z_5 form a triangle, z_1, z_2, z_4 form a triangle, and z_1, z_5 are adjacent. It easy to see now that z_1, z_2, x_3, z_4, z_5 form a copy of K_5.

□

2.6 The diagonal case $\alpha(r,r)$

2.6.1 Proof of Theorem 2.2(iv)

We have seen that the lower bound $\alpha(k,r) \geq k(2r-4)$ in Theorem 2.2(i) is attained for some k. In this subsection, we show that this is not the truth for the diagonal case $k=r\geq 4$, i.e. $\alpha(r,r) \geq r(2r-4)+1$. We shall again use the concept of special vertices introduced in Section 2.5.

Suppose for contradiction that for some $r \geq 4$, $\alpha(r,r) = r(2r-4)$, i.e. there exists a K_r-partite-saturated r-partite graph $G = V_1 \cup V_2 \cup \cdots \cup V_r$ containing an independent set $X = \{x_1, x_2, \ldots, x_r\}$ with $x_i \in V_i$ for all i such that $\sum_{i=1}^{r} d(x_i) = r(2r-4)$. Lemma 2.8 tells us that we must have $d(x_i) = 2r-4$ for all i and each x_i has some neighbor in V_j for $j \neq i$. Therefore, each x_i creates at least two special vertices, and so the sum of the special degrees of the vertices in X^c is at least $2r$. By Lemma 2.18(iv), there is a vertex of special degree at least 3, say $y_1 \in V_1$.

We observe that y_1 has at least two non-neighbors, say $y_2 \in V_2$ and $y_3 \in V_3$ by Lemma 2.18(vi). Since y_1 has special degree at least 3, we can pick $i \geq 4$ such that y_1 is i-special. By Lemma 2.18(v), y_2 and y_3 are neighbors of x_i. Therefore,

$$|N(x_i) \cap N(y_1)| = d(x_i) - |N(x_i) \setminus N(y_1)| \leq (2r-4) - 3 = 2r-7.$$

On the other hand, we shall obtain a contradiction by showing that the graph $H = G[N(x_i) \cap N(y_1)]$ contains at least $\beta_1(r-2, r-2) = 2(r-3)$ vertices. It is sufficient to prove that H is an $(r-2)$-partite K_{r-2}-free graph such that the subgraph induced by any $k-3$ parts contains a K_{r-3}. Clearly, H is K_{r-2}-free since G is K_r-free. The parts of H are $N(x_i) \cap N(y_1) \cap V_j$ for $j \notin [r] \setminus \{1, i\}$. It remains to verify that the deletion of the part $N(x_i) \cap N(y_1) \cap V_j$ does not destroy all the K_{r-3}. Since the non-edge $x_i x_j$ is K_r-saturated in G, the common neighborhood in G of x_i and x_j contains a K_{r-2} consisting of one vertex from each part of $G \setminus \left(V_i \cup V_j\right)$. Since y_1 is i-special, this K_{r-2} must contain y_1, and so the common neighborhood $N(x_i) \cap N(y_1) \cap N(x_j) \subset H$ contains a K_{r-3} not using the vertices of V_j as required. □

2.6.2 Remark on $\beta_2(r, r-1)$

Recall from Proposition 2.7 that $\alpha(r,r) \leq (r-1)\beta_2(r,r-1)$. Thus, a better estimate on β_2 would translate to a better understanding of the saturation numbers. While we could not find the exact value of $\beta_2(r,r-1)$, we suspect that $\beta_2(r,r-1) = 3r-6$ as mentioned in Conjecture 2.13. In this subsection, we make an observation about $\beta_2(r,r-1)$ which can be viewed as a first step towards determining its exact value. For simplicity of notation, let us write $\beta_2(r) = \beta_2(r,r-1)$.

Proposition 2.20. Either

- $\beta_2(r) = 3r-6$ for all $r \geq 3$, or
- $\beta_2(r) \leq (c+o(1))r$ for some constant $c < 3$, as $r \to \infty$.

Proof. The result is an immediate consequence of the following lemma.

Lemma 2.21. $\beta_2(r_1+r_2) \leq \beta_2(r_1) + \beta_2(r_2) + 6$ for $r_1, r_2 \geq 3$.

Proof. For $i \in \{1,2\}$, let $G_i = V_{i,1} \cup V_{i,2} \cup \cdots \cup V_{i,r_i}$ be a K_{r_i-1}-free r_i-partite graph on $\beta_2(r_i)$ vertices such that the subgraph induced by any $r_i - 2$ parts contains a K_{r_i-2}. We shall construct a $K_{r_1+r_2-1}$-free (r_1+r_2)-partite graph G from G_1 and G_2 with $|G| = |G_1| + |G_2| + 6$ by starting with the disjoint union of G_1 and G_2 and then adding six new vertices $U = \{x_1, x_2, y_1, y_2, z_1, z_2\}$ as follows: add x_i, y_i to $V_{i,1}$ and add z_i to $V_{i,2}$ for $i \in \{1,2\}$. Now, join all admissible pairs between U and $V(G) \setminus U$, and add the edges $x_1z_1, x_2z_2, y_1y_2, z_1z_2, y_1z_2, z_1y_2$ inside U.

First, we show that G is $K_{r_1+r_2-1}$-free. Suppose otherwise. Since G_i is K_{r_i-1}-free for $i \in \{1,2\}$, this $K_{r_1+r_2-1}$ must contain at least three vertices forming a triangle in U, contradicting the fact that $G[U]$ is triangle-free. It remains to show that the deletion of any two parts does not destroy all the $K_{r_1+r_2-2}$. Suppose first that both deleted parts are from G_1. Since G_1 contains a K_{r_1-2} not using these two parts and G_2 contains a K_{r_2-2} not using $V_{2,1}$ and $V_{2,2}$, we obtain a $K_{r_1+r_2-2}$ not using the deleted parts, formed by these two cliques and x_2, z_2. Now suppose that one of the deleted parts is from G_1 and the other is from G_2. For $i \in \{1,2\}$, let V_i be a part in $\{V_{i,1}, V_{i,2}\}$ which was not deleted. By construction, $G[U]$ contains an edge between $V_{1,j}$ and $V_{1,l}$ for all $j,l \in \{1,2\}$ and so there exists an edge in $G[U]$ between V_1 and V_2, say e. Since G_1 contains a K_{r_1-2} not using the deleted part in G_1 and V_1, and G_2 contains a K_{r_2-2} not using the deleted part in G_2 and V_2, we obtain a $K_{r_1+r_2-2}$ not using the deleted parts, formed by these two cliques and the endpoints of e. □

Suppose that $\beta_2(s) < 3s-6$ for some $s \geq 3$. We shall show that $\beta_2(r) \leq (c+o(1))r$ with $c = \frac{\beta_2(s)+6}{s} < 3$. Applying the lemma and induction on m, we deduce that $\beta_2(ms) \leq cms - 6$ for all positive integer m. Hence, writing $r = ms+t$ with $3 \leq t \leq s+2$ and applying the lemma again,

$$\beta_2(r) \leq \beta_2(ms) + \beta_2(t) + 6 \leq cms + d \leq \left(c + \frac{d}{r}\right) r = (c+o(1))r$$

where $d = \max\{\beta_2(t) : 3 \leq t \leq s+2\}$. □

2.7 Proof of Theorem 2.3

Theorems 2.1 and Theorem 2.2(ii) imply that

$$sat(n,k,r) = k(2r-4)n + o(n) \text{ if } \begin{cases} k = 2r-3, \text{ or} \\ k \geq 2r-2 \text{ and } r \equiv 0 \mod 2, \text{ or} \\ k \geq 2r-1 \text{ and } r \equiv 2 \mod 3. \end{cases}$$

In this section, we shall show that the $o(n)$ term can be replaced with $O(1)$. The upper bound follows from Proposition 2.5 and Theorem 2.2(ii). We prove that the lower bound holds for any $k \geq r \geq 3$ using the fact that $\beta_1(k-1,r-1) = 2r-4$.

Proposition 2.22. For $k \geq r \geq 3$, there is an integer $C_{k,r}$ such that $sat(n,k,r) \geq k(2r-4)n + C_{k,r}$, for every integer $n \geq 0$.

Proof. Suppose, as we may, that n is sufficiently large. Let $G = V_1 \cup V_2 \cup \cdots \cup V_k$ be a K_r-partite-saturated k-partite graph with $|V_i| = n$ for all i. We shall find a subset U of $V(G)$ of constant size such that every vertex in U^c has at least $2r-4$ neighbors in U. Then we would be done since $e(G) \geq e(U,U^c) \geq (2r-4)(kn-|U|)$. Let v_1 be a vertex of smallest degree in V_1. Having defined $v_1, v_2, \ldots, v_{i-1}$, let $v_i \in V_i$ be a vertex of smallest degree in $V_i \setminus (N(v_1) \cup N(v_2) \cup \cdots \cup N(v_{i-1}))$. We shall take U to be $N(v_1) \cup N(v_2) \cup \cdots \cup N(v_k)$. Now we may assume that $d(v_i) < 2k(2r-4)$ for all $1 \leq i \leq k$. Indeed, if v_i is the first vertex in the sequence such that $d(v_i) \geq 2k(2r-4)$ then we are done since

$$e(G) \geq e(V_i, V_i^c) \geq d(v_i)\left(n - \sum_{j<i} d(v_j)\right) \geq 2k(2r-4)\Big(n - 2k(2r-4)(i-1)\Big) \geq k(2r-4)n$$

for sufficiently large n. Therefore, U has size bounded by a function of k and r. It remains to show that every vertex $v \in U^c$ has at least $2r-4$ neighbors in U. We shall prove that $H = G[N(v) \cap U]$ contains at least $\beta_1(k-1,r-1) = 2r-4$ vertices by showing that H is a K_{r-1}-free $(k-1)$-partite graph such that the subgraph induced by any $k-2$ parts contains a K_{r-2}. Clearly, H is K_{r-1}-free since G is K_r-free. Without loss of generality, $v \in V_1$. The parts of H are $N(v) \cap U \cap V_i$ for $2 \leq i \leq k$. The deletion of the part $N(v) \cap U \cap V_i$ does not destroy all the K_{r-2} since the non-edge vv_i is K_{r-1}-saturated in G, i.e. $N(v) \cap N(v_i) \subset H$ contains a K_{r-2} not using the vertices of V_i. $\square$

2.8 Concluding remarks

We have reduced the problem of determining $sat(n,k,r)$ for large n to that of $\alpha(k,r)$. Although, we have determined $\alpha(k,r)$ for some values of k and r, a large number of cases remain unknown. In particular, the seemingly easiest case when r is fixed and k is large, is still open.

Problem 2.23. Determine $\alpha(k,r)$ for $k \geq 2r-2$ and $r \equiv 1,3 \mod 6$.

For $k \geq 2r-2$ and $r \equiv 0,2,4,5 \mod 6$, we have determined $\alpha(k,r)$ except one missing case when 3 is the smallest divisor of $r-2$ and $k = 2r-2$. Theorem 2.2(i) implies that $\alpha(2r-2,r) \in \{(2r-3)^2, (2r-3)^2-1\}$ and we suspect that $\alpha(2r-2,r) = (2r-3)^2$.

Not only we believe that $\beta_2(k,r) = 4r-k-2$ for $r < k \leq 2r-1$ (see Conjecture 2.13) but we also think that the upper bound $\alpha(k,r) \leq (k-1)\beta_2(k,r-1) \leq (k-1)(4r-k-6)$ in Theorem 2.2(i) is the correct value for $\alpha(k,r)$ in this case.

Conjecture 2.24. $\alpha(k,r) = (k-1)(4r-k-6)$ for $5 \leq r \leq k \leq 2r-4$.

We have shown that $33 \leq \alpha(5,5) \leq 36$. This is the smallest case for which the value of α is not yet known.

Problem 2.25. Find $\alpha(5,5)$.

To prove the lower and upper bounds on $\alpha(k,r)$, we extensively used the bounds on $\beta_1(k,r)$ and $\beta_2(k,r)$. We believe that determining the values of $\beta_i(k,r)$ is an interesting problem on its own.

Problem 2.26. Determine $\beta_i(k,r)$ for $k \geq r \geq 2$ and $2 \leq i \leq k-r+1$.

We end the chapter with a remark on a related problem. Recall that $sat(n,K_r)$ is the minimum number of edges in a K_r-free graph on n vertices but the addition of an edge joining any two non-adjacent vertices creates a K_r. In the pioneer paper of Erdős, Hajnal, and Moon [32], they determined $sat(n, K_r)$ by considering a more general problem where the graphs were not required to be K_r-free. Interestingly, the two problems have the same answer since the extremal graph is K_r-free. We remark that this phenomenon does not happen for partite saturation. Roberts [72] studied the corresponding more general problem for $sat(K_{r\times n}, K_r)$ and showed that the minimum number of edges in a K_r-saturated subgraph of $K_{r\times n}$ where the subgraph is allowed to contain K_r is $\binom{r}{2}(2n-1)$ for $r \geq 4$ and sufficiently large n. On the other hand, Theorem 2.1 and Theorem 2.2 imply that $sat(K_{r\times n}, K_r) \geq r(2r-4)n + o(n) > \binom{r}{2}(2n-1)$ for sufficiently large n.

CHAPTER 3

MAJORITY COLOURINGS OF DIGRAPHS

In this chapter, we solve problems related to a concept called majority coloring recently studied by Kreutzer, Oum, Seymour, van der Zypen and Wood. They raised a problem of determining, for a natural number k, the smallest number $m = m(k)$ such that every digraph can be colored with m colors where each vertex has the same color as at most $1/k$ proportion of its out-neighbors. We show that $m(k) \in \{2k-1, 2k\}$. We also prove a result supporting the conjecture that $m(2) = 3$. Moreover, we prove similar results for a more general concept called majority choosability. This work is joint with António Girão and Teeradej Kittipassorn.

3.1 Results

For a natural number $k \geq 2$, a $\frac{1}{k}$*-majority coloring* of a digraph is a coloring of the vertices such that each vertex receives the same color as at most a $1/k$ proportion of its out-neighbors. We say that a digraph D is $\frac{1}{k}$*-majority m-colourable* if there exists a $\frac{1}{k}$-majority coloring of D using m colors. The following natural question was recently raised by Kreutzer, Oum, Seymour, van der Zypen and Wood [59].

Question 3.1. Given $k \geq 2$, determine the smallest number $m = m(k)$ such that every digraph is $\frac{1}{k}$-majority m-colourable.

In particular, they asked whether $m(k) = O(k)$. Let us first observe that $m(k) \geq 2k-1$. Consider a tournament on $2k-1$ vertices where every vertex has out-degree $k-1$. Any $\frac{1}{k}$-majority coloring of this tournament must be a proper vertex-coloring, and hence it needs at least $2k-1$ colors. Conversely, we prove that $m(k) \leq 2k$.

Theorem 3.2. Every digraph is $\frac{1}{k}$-majority $2k$-colourable for all $k \geq 2$.

This is an immediate consequence of a result of Keith Ball (see [18]) about partitions of matrices. We shall use a slightly more general version proved by Alon [3].

Lemma 3.3. Let $A = (a_{ij})$ be an $n \times n$ real matrix where $a_{ii} = 0$ for all i, $a_{ij} \geq 0$ for all $i \neq j$, and $\sum_j a_{ij} \leq 1$ for all i. Then, for every t and all positive reals $c_1, \ldots, c_t$ whose sum is 1, there is a partition of $\{1, 2, \ldots, n\}$ into pairwise disjoint sets $S_1, S_2, \ldots, S_t$, such that for every r and every $i \in S_r$, we have $\sum_{j \in S_r} a_{ij} \leq 2c_r$.

Proof of Theorem 3.2. Let D be a digraph on n vertices with vertex set $\{v_1, v_2, \ldots, v_n\}$ and write $d^+(v_i)$ for the out-degree of v_i. Let $A = (a_{ij})$ be an $n \times n$ matrix where $a_{ij} = \frac{1}{d^+(v_i)}$ if there is a directed edge from v_i to v_j and $a_{ij} = 0$ otherwise. We apply Lemma 3.3 with $t = 2k$ and $c_i = \frac{1}{2k}$ for $1 \leq i \leq 2k$ obtaining a partition of $\{1, 2, \ldots, n\}$ into sets $S_1, S_2, \ldots, S_{2k}$, such that for every r and every $i \in S_r$, $\sum_{j \in S_r} a_{ij} \leq \frac{1}{k}$. Equivalently, the number of out-neighbors of v_i that have the same color as v_i is at most $\frac{d^+(v_i)}{k}$ where the coloring of D is defined by the partition $S_1 \cup S_2 \cup \cdots \cup S_{2k}$. □

Question 3.1 has now been reduced to whether $m(k)$ is $2k-1$ or $2k$.

Question 3.4. Is every digraph $\frac{1}{k}$-majority $(2k-1)$-colourable?

Surprisingly, this is open even for $k = 2$. Kreutzer, Oum, Seymour, van der Zypen and Wood [59] gave an elegant argument showing that every digraph is $\frac{1}{2}$-majority 4-colourable and they conjectured that $m(2) = 3$.

Conjecture 3.5. Every digraph is $\frac{1}{2}$-majority 3-colourable.

We provide evidence for this conjecture by proving that tournaments are *almost* $\frac{1}{2}$-majority 3-colourable.

Theorem 3.6. Every tournament can be 3-colored in such a way that all but at most 205 vertices receive the same color as at most half of their out-neighbors.

Proof. The proof relies on an observation that in a tournament T, the set $S_i = \{x \in V(T) : 2^{i-1} \leq d^+(x) < 2^i\}$ has size at most 2^{i+1}. Indeed, the sum of the out-degrees of the vertices of S_i is at least $\binom{|S_i|}{2}$, the number of edges inside S_i. On the other hand, this sum is at most $(2^i - 1)|S_i|$ by the definition of S_i. Therefore, $\binom{|S_i|}{2} \leq (2^i - 1)|S_i|$ and hence, $|S_i| \leq 2^{i+1} - 1$.

We proceed by randomly assigning one of three colors to each vertex independently with probability $1/3$. Given a vertex x, let B_x be the number of out-neighbors of x which receive the same color as x. We say that x is *bad* if $B_x > d^+(x)/2$. Trivially $\mathbb{E}(B_x) = d^+(x)/3$, and hence by a Chernoff-type bound, it follows that, for $x \in S_i$,

$$\begin{aligned}\mathbb{P}(x \text{ is bad}) &= \mathbb{P}(B_x > d^+(x)/2) = \mathbb{P}\left(B_x > (1+1/2)\mathbb{E}(B(x))\right) \\ &\leq \exp\left(-\frac{(1/2)^2}{3}\mathbb{E}(B_x)\right) = \exp(-d^+(x)/36) \leq \exp(-2^{i-1}/36).\end{aligned}$$

Notice that if $i \geq 11$ then $\mathbb{P}(x \text{ is bad}) \leq 2^{-(2i-7)}$. Let X denote the total number of bad vertices. Since the vertices of out-degree 0 cannot be bad,

$$\begin{aligned}\mathbb{E}(X) &= \sum_{i\geq 1}\sum_{x\in S_i}\mathbb{P}(x \text{ is bad}) \leq \sum_{i=1}^{10} 2^{i+1}\exp(-2^{i-1}/36) + \sum_{i\geq 11} 2^{i+1}2^{-(2i-7)} \\ &\leq 205 + \sum_{i\geq 11} 2^{-i+8} = 205 + \frac{1}{4} < 206.\end{aligned}$$

Hence, there is a 3-coloring such that all but at most 205 vertices receive the same color as at most half of their out-neighbors. □

Observe also that the same argument proves a special case of Conjecture 3.5.

Theorem 3.7. Every tournament with minimum out-degree at least 2^{10} is $\frac{1}{2}$-majority 3-colorable.

We remark that Theorem 3.6 can be strengthened (205 can be replaced by 7) by solving a linear programming problem. Recall that the expected number of bad vertices of out-degree at least 1024 is at most $1/4$. We shall use linear programming to show that the expected number of bad vertices of out-degree less than 1024 is less than 7.75. Let V_i be the set of vertices of out-degree i for $i \in \{1,2,\ldots,1023\}$ and note that the expected number of bad vertices of out-degree at most 1023 is $f(v_1,v_2,\ldots,v_{1023}) = \sum_{i=1}^{1023} v_i p_i$ where $v_i = |V_i|$ and $p_i = \sum_{j=\lceil \frac{i+1}{2} \rceil}^{i} \binom{i}{j}(1/3)^j(2/3)^{i-j}$. As before, observe that the number of vertices of degree at most i is at most $2i+1$, and therefore, $\sum_{j=1}^{i} v_i \leq 2i+1$, leading to the following linear program.

$$\text{Maximize: } f(v_1,v_2,\ldots,v_{1023})$$

$$\text{Subject to: } \sum_{j=1}^{i} v_j \leq 2i+1, \text{ for } i \in \{1,2,\ldots,1023\}$$

$$\text{Subject to: } v_i \geq 0, \text{ for } i \in \{1,2,\ldots,1023\}.$$

See Appendix 3.2 for the source code. Similarly, we can replace 2^{10} in Theorem 3.7 by 55, by using the same linear program to show that the expected number of bad vertices of out-degree in $[55,1023]$ is less than $3/4$.

Let us now change direction to a more general concept of majority choosability. A digraph is $\frac{1}{k}$*-majority m-choosable* if for any assignment of lists of m colors to the vertices, there exists a $\frac{1}{k}$-majority coloring where each vertex gets a color from its list. In particular, a $\frac{1}{k}$-majority m-choosable digraph is $\frac{1}{k}$-majority m-colourable. Kreutzer, Oum, Seymour, van der Zypen and Wood [59] asked whether there exists a finite number m such that every digraph is $\frac{1}{2}$-majority m-choosable. Anholcer, Bosek and Grytczuk [6] showed that the statement holds with $m = 4$. We generalize their result as follows.

Theorem 3.8. Every digraph is $\frac{1}{k}$-majority $2k$-choosable for all $k \geq 2$.

Theorem 3.8 was independently proved by Fiachra Knox and Robert Šámal [57]. We prove Theorem 3.8 using a slight modification of Lemma 3.3.

Lemma 3.9. Let $A = (a_{ij})$ be an $n \times n$ real matrix where $a_{ii} = 0$ for all i, $a_{ij} \geq 0$ for all

$i \neq j$, and $\sum_j a_{ij} \leq 1$ for all i. Then, for every m and subsets $L_1, L_2, \ldots, L_n \subset \mathbb{N}$ of size m, there is a function $f : \{1,2,\ldots,n\} \to \mathbb{N}$ such that, for every i, $f(i) \in L_i$ and $\sum_{j \in f^{-1}(r)} a_{ij} \leq \frac{2}{m}$ where $r = f(i)$.

Proof. By increasing some of the numbers a_{ij}, if needed, we may assume that $\sum_j a_{ij} = 1$ for all i. We may also assume, by an obvious continuity argument, that $a_{ij} > 0$ for all $i \neq j$. Thus, by the Perron–Frobenius Theorem, 1 is the largest eigenvalue of A with right eigenvector $(1,1,\ldots,1)$ and left eigenvector $(u_1, u_2, \ldots, u_n)$ in which all entries are positive. It follows that $\sum_i u_i a_{ij} = u_j$. Define $b_{ij} = u_i a_{ij}$, then $\sum_i b_{ij} = u_j$ and $\sum_j b_{ij} = u_i \left(\sum_j a_{ij}\right) = u_i$.

Let $f : \{1,2,\ldots,n\} \to \mathbb{N}$ be a function such that $f(i) \in L_i$ and f minimizes the sum $\sum_{r \in \mathbb{N}} \sum_{i,j \in f^{-1}(r)} b_{ij}$. By minimality, the value of the sum will not decrease if we change $f(i)$ from r to l where $l \in L_i$. Therefore, for any $i \in f^{-1}(r)$ and $l \in L_i$, we have

$$\sum_{j \in f^{-1}(r)} (b_{ij} + b_{ji}) \leq \sum_{j \in f^{-1}(l)} (b_{ij} + b_{ji}).$$

Summing over all $l \in L_i$, we conclude that

$$m \sum_{j \in f^{-1}(r)} (b_{ij} + b_{ji}) \leq \sum_{j \in f^{-1}(L_i)} (b_{ij} + b_{ji}) \leq \sum_{j=1}^{n} (b_{ij} + b_{ji}) = 2u_i.$$

Hence, $\sum_{j \in f^{-1}(r)} u_i a_{ij} = \sum_{j \in f^{-1}(r)} b_{ij} \leq \sum_{j \in f^{-1}(r)} (b_{ij} + b_{ji}) \leq \frac{2u_i}{m}$. Dividing by u_i, the desired result follows. □

Proof of Theorem 3.8. The proof is the same as that of Theorem 3.2, using Lemma 3.9 instead of Lemma 3.3. □

In fact, the same statement also holds when the size of the lists is odd.

Corollary 3.10. Every digraph is $\frac{2}{m}$-majority m-choosable for all $m \geq 2$.

This statement generalizes a result of Anholcer, Bosek and Grytczuk [6] where they prove the case $m = 3$ which says that, given a digraph with color lists of size three assigned to the vertices, there is a coloring from these lists such that each vertex has the same color as at most two thirds of its out-neighbors.

We have established that the $\frac{1}{k}$-majority choosability number is either $2k-1$ or $2k$. Let us end this chapter with an analogue of Question 3.4.

Question 3.11. Is every digraph $\frac{1}{k}$-majority $(2k-1)$-choosable?

3.2 Linear program

We use the toolkit [1] to solve the linear program with the following source code:

```
param N := 1024;
param comb 'n choose k' {n in 0..N, k in 0..n} :=
    if k = 0 or k = n then 1 else comb[n-1,k-1] + comb[n-1,k];
param prob 'probability' {n in 0..N} :=
    sum{k in (floor(n/2)+1)..n} comb[n, k] * ((1/3)^k) * ((2/3)^(n-k));

var x{1..N}, integer, >= 0;
subject to constraint{i in 1..N}: sum{j in 1..i} x[j] <= 2*i+1;
maximize expectation: sum{i in 1..N} x[i]*prob[i];

solve;
end;
```

CHAPTER 4

LARGE INDUCED SUBGRAPHS WITH K VERTICES OF ALMOST MAXIMUM DEGREE

In this chapter, we prove that for every integer k, there exist constants $g_1(k)$ and $g_2(k)$ such that the following holds. If G is a graph on n vertices with maximum degree Δ then it contains an induced subgraph H on at least $n - g_1(k)\sqrt{\Delta}$ vertices, such that H contains k vertices of the same degree of order at least $\Delta(H) - g_2(k)$. This solves an approximate version of a conjecture of Caro and Yuster which states that $g_2(k)$ can be taken to be 0 for every k. This work is joint with António Girão.

4.1 Introduction

Given a graph G, let the repetition number, denoted by $rep(G)$, be the maximum multiplicity of a vertex degree. Trivially, any graph G of order at least two contains at least two vertices of the same degree, i.e. $rep(G) \geq 2$. This parameter has been widely studied by several researchers (e.g., [7, 14, 21, 24, 23]), in particular, by Bollobás and Scott, who showed that for every $k \geq 2$ there exist triangle-free graphs on n vertices with $rep(G) \leq k$ for which $\alpha(G) = (1 + o(1))n/k$ ([14]). As there are infinitely many graphs having repetition number two, it is natural to ask what is the smallest number of vertices one needs to delete from a graph in order to increase the repetition number of the remaining induced subgraph. This question was partially answered by Caro, Shapira and Yuster in [20], indeed, they proved that for every k there exists a constant $C(k)$ such that given any graph on n vertices one needs to remove at most $C(k)$ vertices and thus obtain an induced subgraph with at least $\min\{k, n - C(k)\}$ vertices of the same degree. Related to this question, Caro and Yuster ([22]) considered the problem of finding the largest induced subgraph H of a graph G which contains at least k vertices of degree $\Delta(H)$. To do so they defined $f_k(G)$ to be the smallest number of vertices one needs to remove from a graph G

such that the remaining induced subgraph has its maximum degree attained by at least k vertices. They found examples of graphs on n vertices for which $f_2(G) \geq (1-o(1))\sqrt{n}$ and conjectured $f_k(G) \leq O(\sqrt{n})$ for every graph G on n vertices. In the same paper they established the conjecture for $k \leq 3$.

The following more general conjecture was posed recently by Caro, Lauri and Zarb in [19].

Conjecture 4.1. For every $k \geq 2$ there is a constant $g(k)$ such that given a graph G with maximum degree Δ, one can remove at most $g(k)\sqrt{\Delta}$ vertices such that the remaining subgraph $H \subseteq G$ has at least k vertices of degree $\Delta(H)$.

Let us define $g(k,\Delta) = \max\{f_k(G) : \Delta(G) \leq \Delta\}$. In the same paper, they proved that $g(2,\Delta) = \lceil \frac{3+\sqrt{8\Delta+1}}{2} \rceil$ and stated that $g(3,\Delta) \leq 42\sqrt{\Delta}$. We should point out that, if true, the conjecture is best possible, as there are graphs on n vertices found in [19] for which any induced subgraph on more than $n - \frac{k}{2}\sqrt{\Delta}$ does not contain k vertices of the same maximum degree. We shall present such constructions in Section 4.3.

In this chapter we prove the following approximate version of Conjecture 4.1

Theorem 4.2. For every positive integer k, there exist constants $g_1(k)$ and $g_2(k)$ such that the following holds. If G is a graph on n vertices with maximum degree Δ then it contains an induced subgraph H on at least $n - g_1(k)\sqrt{\Delta}$ vertices, such that H has k vertices of the same degree at least $\Delta(H) - g_2(k)$.

4.2 Proofs

First, we shall introduce the following definitions. Let n be an integer and $A_1 \cup A_2 \cup \ldots \cup A_t$ be a partition of the set $\{1,2,\ldots,n\}$ into t sets. Moreover, let $r_1 > r_2 > r_3 > \ldots > r_t$ be a strictly decreasing sequence of non-negative integers. We shall say that a multiset $\mathscr{A}$ consisting of subsets of $[n]$ is an $(r_1, r_2, \ldots, r_t)$-uniform cover of

$\{1,2\ldots,n\}$ if for every $i \in \{1,\ldots,t\}$ and $j \in A_i$, we have $|\{A \in \mathscr{A} : j \in A\}| = r_i$. Note that in a multiset we allow repetitions.

We call an $(r_1,r_2,\ldots,r_t)$-uniform cover $\mathscr{A}$ of $\{1,2,\ldots,n\} = A_1 \cup A_2 \cup \ldots \cup A_t$ *irreducible* if there is no proper $(r'_1,\ldots,r'_t)$-uniform cover $\mathscr{B} \subset \mathscr{A}$, for some strictly decreasing sequence of non-negative integers $r'_1 > r'_2 > \ldots > r'_t$.

Given a uniform cover $\mathscr{A}$ of $\{1,2,\ldots,n\}$ and a subset $B \subseteq \{1,2,\ldots,n\}$ we define $w_{\mathscr{A}}(B)$ to be the number of times B appears in $\mathscr{A}$.

Lemma 4.3. For all $n \in \mathbb{N}$, there exists $f(n)$ such that for any $1 \leq t \leq n$ and any partition of $\{1,2,\ldots,n\}$ into t sets $A_1,A_2,\ldots,A_t$, every $(r_1,r_2,\ldots,r_t)$-uniform cover $\mathscr{A}$ of $\{1,2,\ldots,n\}$ contains a $(r'_1,r'_2,\ldots,r'_t)$-uniform sub-cover $\mathscr{B} \subset \mathscr{A}$ with $r'_1 \leq f(n)$.

Proof. We shall prove there are only finitely many *irreducible* covers. For otherwise, let us assume there exists an infinite sequence $\{B_i\}_{i\in\mathbb{N}}$ of *irreducible* uniform covers. Since there are only finitely many partitions of $\{1,2,\ldots,n\}$, we may pass to an infinite subsequence $\{B_{l_i}\}_{i\in\mathbb{N}}$ of uniform covers of the same partition of $\{1,2,\ldots,n\}$. Now, choose $A \subseteq \{1,2,\ldots,n\}$ and consider the sequence of non-negative integers $\{w_{B_{l_i}}(A)\}_{i\in\mathbb{N}}$, clearly it must contain an infinite non-decreasing subsequence $w_{B_{l_{i_1}}}(A) \leq w_{B_{l_{i_2}}}(A) \leq \ldots$. We restrict our attention to this subsequence of the uniform covers $B_{l_{i_1}}, B_{l_{i_2}}, \ldots$ and iteratively apply the same argument for the remaining subsets of $\{1,2,\ldots,n\}$, always passing to a subsequence of the previous sequence of uniform covers. After we have done it for every subset of $\{1,2,\ldots,n\}$, we must end up with two distinct *irreducible* uniform covers (actually an infinite sequence) $\mathscr{A},\mathscr{B}$ for which $w_{\mathscr{A}}(F) \leq w_{\mathscr{B}}(F)$ for every $F \subseteq \{1,2,\ldots,n\}$. This implies $\mathscr{A} \subseteq \mathscr{B}$, which is a contradiction. Take $f(n)$ to be the maximum r_1 over all *irreducible* uniform covers of $\{1,2,\ldots,n\}$. □

Lemma 4.4. For every $n \in \mathbb{N}$, there exists $f(n)$ such that the following holds. Let $G = (A,B)$ be a bipartite graph with $A = \{x_1,x_2,\ldots,x_n\}$. Then there exists a subset $W \subseteq V(B)$ of size at most $n \cdot f(n) = f'(n)$, such that the induced bipartite graph

$G' = G[A,(B \setminus W)]$ has the property that

$$\text{if } d_G(x_i) > d_G(x_j) \text{ then } d_G(x_i) - d_{G'}(x_i) > d_G(x_j) - d_{G'}(x_j).$$

Proof. Partition A into $A_1,\ldots,A_t$, so that two vertices belong to the same part if they have the same degree. Let r_i be the degree of the vertices in A_i. We may assume that $r_1 > r_2 > \cdots > r_t$. The lemma follows as a corollary of Lemma 4.3. Indeed, for every vertex $w \in B$, let $A_w \subseteq \{1,2,\ldots,n\}$ such that $i \in A_w$ if x_i is a neighbor of w in G. Note that $\mathscr{A} = \{A_w : w \in B\}$ is an $(r_1,r_2,\ldots,r_t)$-uniform cover of $\{1,2,\ldots,n\}$. Applying now Lemma 4.3, we can find a $(r_1',r_2',\ldots,r_t')$-uniform sub-cover $\mathscr{B} \subseteq \mathscr{A}$ with $r_1' \le f(n)$. Let $W = \{w \in B : A_w \in \mathscr{B}\}$ and $G' = G[A,(B \setminus W)]$. It is easy to see that $|W| \le n \cdot f(n)$ and that the property is satisfied by the definition of uniform cover. □

Given a positive integer k and a graph G with the vertex set $\{x_1,\ldots,x_n\}$ such that $d(x_1) \ge \cdots \ge d(x_n)$, let $r_k(G) := \Delta(G) - d_G(x_k)$ be the difference between the maximum degree and the degree of vertex x_k.

Theorem 4.5. For every positive integer k there exists $h(k)$ such that the following holds. If G is a graph on n vertices with maximum degree Δ then it contains an induced subgraph H on at least $n-(h(k)+k)\sqrt{\Delta}$ vertices, such that $r_k(H) \le h(k)\cdot k$.

Proof. The proof consists of two parts. Firstly, we shall show that we can remove at most $k\sqrt{\Delta}$ vertices from G so that in the remaining graph H' we have $r_k(H') \le \sqrt{\Delta}$. Then we iteratively apply Lemma 4.4 (at most $\sqrt{\Delta}$ times) in order to obtain an induced subgraph H of H' on at least $n-(h(k)+k)\sqrt{\Delta}$ vertices such that $r_k(H) \le h(k)\cdot k$. We may take $h(k)$ to be $f'(k)$ from Lemma 4.4. We start with the first part of the proof.

Claim 4.6. There is an induced subgraph H' of G on at least $n-k\sqrt{\Delta}$ vertices such that $r_k(H') \le \sqrt{\Delta}$.

Proof of Claim 4.6. The idea of the proof is to keep removing some k vertices of highest possible degrees and observe that the maximum degree on the induced remaining graph must have decreased considerably. Indeed, consider the following procedure. Let $G_0 = G$ and suppose that $G_0 \supset \cdots \supset G_i$ have been defined. If G_i does not have the required property then, let G_{i+1} be obtained from G_i by removing some k vertices with largest degrees in G_i. Notice that $\Delta(G_{i+1}) \leq \Delta(G_i) - \sqrt{\Delta}$ since, by assumption, there were at most k vertices in G_i having degrees in the range $[\Delta(G_i), \Delta(G_i) - \sqrt{\Delta}]$. Also $|G_{i+1}| = |G_i| - k$. Observe that the procedure will stop after at most $\sqrt{\Delta}$ steps, as otherwise the obtained graph would have maximum degree 0. Since $|G_i| \geq n - i \cdot k$ we have that $|H'| \geq n - k\sqrt{\Delta}$. □

We now proceed to the second part of the proof and iteratively apply Lemma 4.4. In each step we remove at most $h(k)$ vertices from H' while decreasing the value of r_k and we stop when r_k is at most $k \cdot h(k)$.

Let $H_0 = H'$ and suppose that $H_0, \ldots, H_i$ have already been defined. If $r_k(H_i) \leq k \cdot h(k)$ then we are done, so we may assume that $r_k(H_i) > k \cdot h(k)$. Let $A = \{x_1, \ldots, x_k\}$ be a set of k vertices with the largest degrees in H_i and write B for $H_i \setminus A$. Without loss of generality we may assume that $d_{H_i}(x_1) \geq \cdots \geq d_{H_i}(x_k)$. Since $r_k(H_i) > k \cdot h(k)$ there must exist $l \in \{2, \ldots, k\}$ such that $d_{H_i}(x_l) > d_{H_i}(x_{l-1}) + h(k)$. Now consider the bipartite subgraph $K = H_i[A, B]$. By Lemma 4.4, with $G = K$ and $n = k$, we can remove a set $W \subset B$ of at most $f'(k) = h(k)$ vertices from B, and obtain $K' = H_i[A, (B \setminus W)]$ such that

$$\text{for any } x, y \in A, \text{ if } d_K(x) < d_K(y) \text{ then } d_K(x) - d_{K'}(x) < d_K(y) - d_{K'}(y). \tag{4.1}$$

Let $H_{i+1} = H_i \setminus W$ (hence $|H_{i+1}| \geq |H_i| - |W| \geq |H_i| - h(k)$). The following claim asserts that the above procedure will stop after at most $\sqrt{\Delta}$ steps.

Claim 4.7. $r_k(H_{i+1}) < r_k(H_i)$.

Proof of Claim 4.7. Let z be a vertex with the maximum degree and w a vertex with the

k'th largest degree in H_{i+1}. Observe that $z = x_t$ for some $t \geq l$ and $d_{H_{i+1}}(w) \geq d_{H_{i+1}}(x_s)$ for some $s < l$. First, notice that $d_{H_i}(x_t) - d_{H_i}(x_s) \leq d_{H_i}(x_1) - d_{H_i}(x_k) = r_k(H_i)$. Hence, $r_k(H_{i+1}) = d_{H_{i+1}}(z) - d_{H_{i+1}}(w) \leq d_{H_{i+1}}(x_t) - d_{H_{i+1}}(x_s) < d_{H_i}(x_t) - d_{H_i}(x_s) \leq r_k(H_i)$, where the strict inequality follows from (4.1) since $d_K(x_t) > d_K(x_s)$. □

As in each iteration the value of r_k decreases, we must stop after at most $r_k(H') = \sqrt{\Delta}$ steps thus getting an induced subgraph $H \subset H'$ with $r_k(H) \leq k \cdot h(k)$ and $|H| \geq |H'| - h(k)\sqrt{\Delta} \geq n - (h(k) + k)\sqrt{\Delta}$. □

In order to prove Theorem 4.2 we need the following theorem of Caro, Shapira and Yuster, appearing in [20], whose proof is inspired by the one used by Alon and Berman in [4].

Theorem 4.8. For positive integers r, d, q, the following holds. Any sequence of $n \geq (\lceil q/r \rceil + 2)(2rd + 1)^d$ elements of $[-r, r]^d$ whose sum, denoted by z, is in $[-q, q]^d$ contains a subsequence of length at most $(\lceil q/r \rceil + 2)(2rd + 1)^d$ whose sum is z.

As usual, we write $R(k)$ (see e.g. [13]) for the *two colored Ramsey number*, the least integer n such that in any two coloring of the edges of the complete graph on n vertices, there is a monochromatic K_k.

Proof of Theorem 4.2. Firstly, we apply Theorem 4.5 with $k = R(k)$ to find a large induced subgraph $G' \subset G$ of order at least $n' \geq n - (h(R(k)) + R(k))\sqrt{\Delta}$ and with vertex set $\{x_1, \ldots, x_{n'}\}$ where $d(x_1) \geq d(x_2) \geq \cdots \geq d(x_{n'})$ and $d(x_1) - d(x_{R(k)}) \leq h(R(k)) \cdot R(k) = M$. Now we follow the proof of Theorem 1.1 in [20].

By the definition of $R(k)$ we can find a set S of k vertices in $\left\{x_1, \ldots, x_{R(k)}\right\}$ that induces either a complete graph or an independent set.

Without loss of generality, assume that $S = \{v_{n'-k+1}, \ldots, v_{n'}\}$ and $V(G) \setminus S = \{v_1, \ldots, v_{n'-k}\}$. Let $e(v_i, v_j)$ be equal to 1 if there is an edge between v_i and v_j, and 0 otherwise. We construct a sequence X of $n' - k$ vectors $w_1, \ldots, w_{n'-k}$ in $[-1, 1]^{k-1}$

as follows. The coordinate j of w_i is $e(v_{n'-k+j}, v_i) - e(v_{n'}, v_i)$ for $i = 1, \ldots, n'-k$ and $j = 1, \ldots, k-1$. It is clear that $e(v_{n'-k+j}, v_i) - e(v_{n'}, v_i) \in [-1, 1]$ as required. Consider the sum of all the j'th coordinates,

$$\begin{aligned}\sum_{i=1}^{n'-k} \left(e(v_{n'-k+j}, v_i) - e(v_{n'}, v_i)\right) &= \sum_{i=1}^{n'-k} e(v_{n'-k+j}, v_i) - \sum_{i=1}^{n'-k} e(v_{n'}, v_i) \\ &= (d(v_{n'-k+j}) - a) - (d(v_{n'}) - a) \\ &= d(v_{n'-k+j}) - d(v_{n'}) \leq M,\end{aligned}$$

where $a = k-1$ if $G'[S]$ is complete, and $a = 0$ otherwise. Hence,

$$z = \sum_{i=1}^{n'-k} w_i \in [-M, M]^{k-1}.$$

By Theorem 4.8, with $d = k-1$ and $q = M$, there is a subsequence of X of size at most $(M+2)(2k-1)^{k-1}$ whose sum is z. Deleting the vertices of G' corresponding to the elements of this subsequence results in an induced subgraph $H \subset G'$ in which all the k vertices of S have the same degree of order at least $\Delta(H) - \left(M + (M+2)(2k-1)^{k-1}\right)$. Choosing $g_1(k) = g_2(k) = h(R(k))(4k)^k$ we conclude the theorem. □

4.3 Remarks

In the previous section, we proved that every graph contains a large induced subgraph with at least k vertices having the same degree of order almost the maximum degree. Note that Theorem 4.2 is sharp up to the size of the functions $g_1(k)$ and $g_2(k)$. Indeed, there are graphs for which one needs to remove ”roughly” $\frac{k}{2}\sqrt{\Delta}$ vertices to force the remaining subgraph to have k vertices with the same degree "near" the maximum degree. For any k and Δ, let G^{Δ} be the disjoint union of the stars $K_{1,n_1}, \ldots, K_{1,n_t}$, where $n_i = i \cdot \sqrt{\Delta}$, for $i \in \left\{1, \ldots, t = \sqrt{\Delta}\right\}$ and let G_k^{Δ} to be the disjoint union of $k/2$ copies of G^{Δ}. It is easy to see that, for any constant D, one needs to remove at least $\frac{k}{2}\sqrt{\Delta} - \frac{k}{2}D$ vertices from G_k^{Δ} in order to obtain an induced graph H with k vertices of the same degree of order at least $\Delta(H) - D$.

Whether removing $C(k)\sqrt{\Delta}$ vertices is enough to force the remaining induced subgraph to have at least k vertices of exactly maximum degree remains an interesting open question.

CHAPTER 5

BOUNDS ON THE GRAPH BURNING NUMBER

In this chapter we prove a few results concerning the burning number, $b(G)$, of a graph G which is a graph parameter defined by Bonato, Janssen, and Roshanbin [16] which, supposedly, measures the speed of the spread of contagion in a graph. We show that for any connected graph G on n vertices, its burning number is bounded above by $\left\lceil \sqrt{\frac{4}{3}n} \right\rceil$. This makes a progress towards a conjecture of Bonato, Janssen, and Roshanbin who conjectured that any connected graph burns in at most $\sqrt{n}$ rounds. Moreover, we prove that if G is a disjoint union of k paths then $b(G) \leq \left\lceil \sqrt{|G| + (k-1)^2} \right\rceil$, which we later use to show that $b(S) \leq \lceil \sqrt{n} \rceil$, for any spider graph S on n vertices. This work is joint with Kazuhiro Nomoto, Julian Sahasrabudhe, and Richard Snyder.

5.1 Introduction

Graph burning is a deterministic process defined on a graph, which was introduced by Bonato, Janssen, and Roshanbin [16], and which is supposed to model the expansion of a fire in a graph. In each step, first the fire spreads from burning vertices to their neighbors that are not already burning, then a new fire starts at some, not yet burning, vertex. The *burning number* of a graph G, denoted by $b(G)$, is the smallest possible number of steps needed to burn the whole graph G. The process was inspired by other graph processes like firefighting, graph cleaning, bootstrap percolation (see for example [38, 5, 8]).

For a vertex $v \in G$ and a non-negative integer r, define, $B_G(v,r)$ to be the set of vertices in G which are at distance at most r from v. For brevity we shall drop the subscript and simply write $B(v,r)$, when there is no risk of confusion.

It is easy to see that the problem of determining the burning number of a graph G is equivalent to finding the smallest integer k such that one can cover the vertices of G with some graphs $H_0, \dots, H_{k-1}$ such that H_i has radius i. Alternatively, we can define $b(G)$ to be the smallest integer k such that there is a sequence of vertices $x_0, \cdots, x_{k-1}$ in G such that

$$V(G) = B(x_0, 0) \cup B(x_1, 1) \cup \cdots \cup B(x_{k-1}, k-1),$$

or, equivalently, for every $y \in G$ there is $i \in \{0, \cdots, k-1\}$ such that $d(x_i, y) \leq i$.

In their original paper, Bonato, Janssen, and Roshanbin asked the following conjecture, which attracted considerable attention.

Conjecture 5.1 (Bonato, Janssen, Roshanbin [16])**.** Let G be a connected graph on n vertices. Then

$$b(G) \leq \left\lceil \sqrt{n} \right\rceil.$$

If the conjecture is true then the result is best possible, as seen by considering a path on n vertices. The conjecture remains open but, nevertheless, certain progress towards it has been made. In the original paper [16] the authors proved that for any connected graph G on n vertices we have $b(G) \leq 2\sqrt{n} - 1$. This was later improved by Bessy, Bonato, Janssen, Rautenbach, and Roshanbin [9] who showed that $b(G) \leq \sqrt{\frac{32}{19}\frac{n}{1-\varepsilon}} + \sqrt{\frac{27}{19\varepsilon}}$ and $b(G) \leq \sqrt{\frac{12n}{7}} + 3 \sim 1.309\sqrt{n}$, for every $\varepsilon \in (0,1)$. Finally, Land and Lu [61] showed the bound $b(G) \leq \left\lceil \frac{-3+\sqrt{24n+33}}{4} \right\rceil \sim 1.22\sqrt{n}$. We make a further improvement and show the following.

Theorem 5.2. Let G be a connected graph on n vertices. Then

$$b(G) \leq \left\lceil \sqrt{\frac{4}{3}n} \right\rceil \sim 1.15\sqrt{n}.$$

The burning number has been also studied for other classes of graphs. Mitsche, Prałat, and Roshanbin [68] considered random binomial graphs, random geometric graphs, and the Cartesian products of paths. Fitzpatrick and Wilm [39] studied the graph burning numbers of circulant graphs. Sim, Tan, Wong [74] gave asymptotically tight bounds for the class of generalized Petersen graphs. Bonato, and Lidbetter [17] proved the following two bounds on the burning number of (disjoint) union of paths.

Theorem 5.3 (Bonato, Lindbetter [17])**.** Let G be a union of k paths on n vertices in total. Then $b(G) \leq \lfloor \frac{n}{2k} \rfloor + k$. Moreover, when $k \leq \lceil \sqrt{n} \rceil$ then $b(G) \leq \lceil \sqrt{n} + \frac{k-1}{2} \rceil$.

They also showed that spider graphs (trees with exactly one vertex of degree at least 3) burn in $\lceil \sqrt{n} \rceil$ rounds. Here we obtain better than in Theorem 5.3 bounds on the burning numbers of unions of paths.

Theorem 5.4. Let G be a union of k paths on n vertices in total. Then

$$b(G) \leq \left\lceil \sqrt{n + (k-1)^2} \right\rceil .$$

Observe that when $k \geq \frac{n+1}{2}$, then Theorem 5.4 implies that $b(G) \leq k$, which together with the trivial lower bound $b(G) \geq k$, gives $b(G) = k$. Note also that the above theorem is tight for every n and $k \leq \frac{n}{2}$ as well, since it takes $\left\lceil \sqrt{n + (k-1)^2} \right\rceil$ steps to burn a union of $k-1$ paths of order 2 and one path of order $n - 2k + 2$. Consequently, we give an alternative proof of the fact that $b(G) \leq \lceil \sqrt{n} \rceil$ when G is a spider graph.

Theorem 5.5. Let G be a spider on n vertices. Then

$$b(G) \leq \lceil \sqrt{n} \rceil .$$

5.2 Tight bounds on burning numbers of spiders

A *spider* graph is a tree in which exactly one vertex has degree at least three and all the other vertices have degrees at most two. We shall call the unique vertex of degree at least 3 the *head* of the spider. The paths obtained by removing the head from the spider graph shall called the *legs* of the spider graph. The aim of this section is to prove Theorem 5.5 and Theorem 5.4. Theorem 5.4 is a corollary of the following lemma.

Lemma 5.6. Suppose $n = n_1 + \cdots + n_k$ where n_i is a positive integer, for every $i \in [k]$. Then for any positive integer t such that $t^2 \geq n + (k-1)^2$ there is a partition of integers $\{0, \cdots, t-1\}$ into $A_1, \cdots, A_k$ such that, for every $i \in [k]$

$$\sum_{a \in A_i} (2a+1) \geq n_i.$$

Let us first deduce Theorem 5.4 from Lemma 5.6.

Proof of Theorem 5.4. Let G be a union of k paths $P_1, P_2, \ldots, P_k$. Write n_i for $|P_i|$ and $|G| =: n = n_1 + n_2 + \ldots n_k$. By Lemma 5.6, for $t = \left\lceil \sqrt{n + (k-1)^2} \right\rceil$, there is a partition of the integers $\{0, \ldots, t-1\}$ into sets $A_1, \ldots, A_k$ such that for every $i \in \{1, \ldots, k\}$, we have $\sum_{a \in A_i} (2a+1) \geq n_i$. It is easy to see that it is possible to cover the vertices of P_i using balls of radii in A_i. □

Proof of Lemma 5.6. Given a set S let $f(S) = \sum_{s \in S} (2s+1)$. Let $\mathscr{A} = \{A_1, \cdots, A_k\}$ be a partition of $\{0, \cdots, t-1\}$ minimizing the function

$$F(\mathscr{A}) = \sum_{i=1}^{k} \left(f(A_i) - n_i\right)^2, \tag{5.1}$$

subject to every element of the partition being non-empty, i.e., $A_i \neq \emptyset$, for every i. We shall show that this partition satisfies the conclusion of the lemma. Assume for the sake of contradiction that the partition $\mathscr{A}$ does not satisfy the conclusion of the lemma, which means that for some i we have $f(A_i) - n_i < 0$.

For $p \in \{0, \cdots, t-1\}$ let $A^{(p)} = A_i$ and $n^{(p)} = n_i$, for i such that $p \in A_i$. Now, define $g(p) = f(A^{(p)}) - n^{(p)}$. We claim that $g(p)$ does not grow too fast as a function of p.

Claim 5.7. For any $p \in \{0, \cdots, t-2\}$, $g(p+1) \leq g(p) + 2$.

Proof. Suppose for contradiction that there is $p \in \{0, \cdots, t-2\}$ with $g(p+1) \geq g(p) + 3$. We shall construct another partition with a smaller square sum 5.1, contradicting the minimality of $\mathscr{A}$. For brevity, let us write $B = A^{(p)}$ and $C = A^{(p+1)}$. Let $\mathscr{A}' = \{A'_1, \cdots, A'_k\}$ where

$$A'_i = \begin{cases} (A_i \setminus \{p\}) \cup \{p+1\}, & \text{if } A_i = B \\ (A_i \setminus \{p+1\}) \cup \{p\}, & \text{if } A_i = C \\ A_i, & \text{otherwise.} \end{cases}$$

It is easy to check that $F(\mathscr{A}) - F(\mathscr{A}') = 4(g(p+1) - g(p) - 2) \geq 4(3-2) = 4 > 0$. □

Our second claim says that 0 belongs to a set A_i such that $f(A_i) - n_i \leq 0$, otherwise we could remove 0 from A_i and put in a set A_j, for some j such that $f(A_j) - n_j < 0$, decreasing the square sum and obtaining a contradiction.

Claim 5.8. $g(0) \leq 0$.

Proof. Assume for contradiction that $g(0) > 0$. Suppose $0 \in A_i$ and let j be any integer such that $A_j - n_j < 0$. Observe that since n_i is a positive integer and $f(A_i) > n_i$, and $0 \in A_i$, it follows that $|A_i| \geq 2$. We can therefore move 0 from A_i to A_j which will result in a smaller square sum 5.1. This contradicts the minimality of $\mathscr{A}$. □

Without loss of generality we can assume that $f(A_i) - n_i \leq f(A_j) - n_j$ for $i < j$. Note that $f(A_1) - n_1 < 0$. It follows from the above two claims that $f(A_i) - n_i \leq 2i - 3$. Therefore

$$t^2 = \sum_{i=1}^{k} f(A_i) = n + \sum_{i=1}^{k} (f(A_i) - n_i) \leq n + \sum_{i=1}^{k} (2i-3) = n + (k-1)^2 - 1,$$

contradicting the assumption that $t^2 \geq n + (k-1)^2$. □

Now we are ready to deduce Theorem 5.5 from Theorem 5.4.

Proof of Theorem 5.5. Let G be a spider with the head v, on $n = k^2 + \ell$ vertices, where $1 \leq \ell \leq 2k+1$. We shall show, using induction on n, that G can be burned in at most $k+1$

steps. The base case, when $n \leq 4$, is trivial as the star $K_{1,3}$ is the only spider on fewer than 5 vertices, in which case G burns in two steps. Suppose that the theorem holds for every integer n' such that $4 \leq n' < n$.

Observe that when G has a leg of *length* (where *length* of a leg is the number of vertices on the leg, not counting the head of the spider) at least ℓ, then we can cover ℓ vertices of the leg with a ball of radius k, obtaining a spider or a path on k^2 vertices, which by induction burns in at most k steps. Hence, in total, G burns in at most $k+1$ steps. We can therefore assume that every leg of G has length at most $\ell - 1 \leq 2k$.

Let t be the number of legs of length at least $k+1$ in G. It follows from an easy calculation that $t \leq k$. We shall consider two cases, first when $t < k$ and second when $t = k$. Suppose first that $t < k$.

Consider the ball $B = B(v,k)$ and let $H = G \setminus B$. Let w be the total number of vertices on legs of length at most k in G. Observe that any leg in G of length at most k, is completely covered by the ball B. Therefore H is a union of t paths, with total number of vertices $n - tk - 1 - w \leq k^2 + 2k - tk - w$. On the other hand, by the assumption that every leg has fewer than $l \leq 2k+1$ vertices, we have that H has at most $t(l-1-k) \leq tk$ vertices. Therefore, combining these two bounds we obtain that

$$|H| \leq \min\left\{tk, k^2 + 2k - tk - w\right\}.$$

Applying Theorem 5.4, we see that

$$b(H) \leq \left\lceil \sqrt{\min\{tk, k^2 + 2k - tk - w\} + (t-1)^2} \right\rceil.$$

Therefore as long as $f_k(t) = \min\left\{tk, k^2 + 2k - tk - w\right\} + (t-1)^2$ does not exceed k^2 we are done. Easy calculation shows that $f_k(1) \leq k \leq k^2$, for $k \geq 2$, and $f_k(2) \leq 2k+1 \leq k^2$, for $k \geq 3$. For $3 \leq t \leq k-1$ and $k \geq 4$ we have

$$\begin{aligned} f_k(t) \leq g_k(t) &= k^2 + 2k - tk + (t-1)^2 \leq g_k(3) = g_k(k-1) \\ &= k^2 - k + 4 \leq k^2. \end{aligned}$$

Therefore we can assume that $t = k$. When $w \geq 1$ then

$$f_k(k) \leq 2k - w + (k-1)^2 = k^2 + 1 - w \leq k^2.$$

We can hence assume that no leg has length smaller than $k+1$. We shall consider two cases depending on the distribution of the lengths of the legs of G.

Case 1 **- there is a leg of length exactly** $k+1$. Place a ball of radius k at a vertex on a leg of length exactly $k+1$, at distance 1 from the head of G. The remaining, uncovered, graph H consists of $k-1$ paths with total number vertices at most
$n-(k+1+1+(k-1)(k-1)) = n-k^2+k-3 \leq (k+1)^2-k^2+k-3 \leq 3k-2$. By Theorem 5.4 we have $b(H) \leq \left\lceil \sqrt{3k-2+(k-2)^2} \right\rceil = \left\lceil \sqrt{k^2-k+2} \right\rceil \leq k$ (note that here we need the assumption that $k \geq 2$), hence we are done.

Case 2 **- all legs have length** $k+2$. Place a ball of radius k at any vertex at distance exactly 2 from the head of the spider G. It is easy to see that the remaining graph consists of $k-1$ paths of length exactly 4. We can cover one of the paths using balls of radius 0 and 1, and the remaining $k-2$ paths by the balls of radii $2, \ldots, k-1$.

It is easy to check that there are no other cases. □

5.3 General Bound on the Burning Numbers of Connected Graphs

In this section we prove Theorem 5.2, i.e., that any connected graph on n vertices can be burned in $\left\lceil \sqrt{\frac{4}{3}n} \right\rceil$ rounds. It is clear that it is enough to verify the bound for trees.

Let T be a tree of diameter d and let $P = \{x_1, \cdots, x_{d+1}\}$ be a longest path in T. For each $x_i \in P$ let T_i be the tree rooted at x_i consisting of x_i and all the vertices which can be reached from x_i not using any other vertex of P. It is easy to see that for any i the height of T_i is strictly less than i, as otherwise there would be a path in T strictly longer than P, contradicting the maximality of P. We say that a ball $B \subset V(T)$ covers A *nicely* if $A \subseteq B$ and $T \backslash A$ remains connected.

The following, rather technical lemma, is the heart of the proof.

Lemma 5.9. Fix a non-negative integer a. Let $X = \{r_0, \cdots, r_s\}$ with $a \leq r_i < r_{i+1} < \frac{d-1}{2}$ be a set of non-negative integers. Assume that for every $r_i \in X$ there is no set of at least $2r_i + 1 - a$ vertices which can be covered nicely with a ball of radius r_i. Let $B_i = B(x_{r_i+1}, r_i)$ be the (closed) ball of radius r_i centered at x_{1+r_i} and let j_i be the smallest non-negative integer such that $T_{j_i} \not\subseteq B_i$. Then the following is true for any $t \in \{0, \ldots, s\}$.

1. $r_t + 1 < j_t \leq 2r_t + 2 - 2t - a$ (and hence $r_t \geq 2t + a$),

2. $\sum_{i=1}^{j_t} |T_i \cap B_t| \geq 2t + a + j_t$.

Proof. We will use induction on t. The first inequality in the first part holds for any t trivially because T_i has height strictly less than i, for any i, hence $T_i \subseteq B_t$, for every $i \leq r_t + 1$. Let us consider the base case when $t = 0$. It is easy to show that $j_0 \leq 2r_0 + 1 - a$, as otherwise we obtain a contradiction by observing that B_0 covers nicely at least $2r_0 + 1 - a$ vertices. To prove the second part, suppose $j_0 = r_0 + 1 + b$, for some non-negative integer b. We have that T_{j_0} has height bigger than $r_0 - b$ as otherwise T_{j_0} would be covered by B_0. Hence

$$\sum_{i=1}^{j_0} |T_i \cap B_0| \geq r_0 + 1 + b + r_0 - b = 2r_0 + 1 \geq j_0 + a.$$

Now suppose the statement of the lemma is true for some non-negative integer $t < s$. We shall show it is true for $t + 1$. Assume first that the first part does not hold, hence

$$j_{t+1} > 2r_{t+1} + 2 - 2(t+1) - a = 2r_{t+1} - 2t - a \geq 2r_t + 2 - 2t - a \geq j_t,$$

where the last inequality follows by the induction hypothesis. It means that B_{t+1} covers T_i nicely, for every $i \leq 2r_{t+1} - 2t - a$, hence the number of nicely covered vertices is at least

$$\begin{aligned}\sum_{i=1}^{2r_{t+1}-2t-a} |T_i| &\geq \sum_{i=1}^{j_t} |T_i| + 2r_{t+1} - 2t - a - j_t \\ &\geq \sum_{i=1}^{j_t} |T_i \cap B_t| + 1 + 2r_{t+1} - 2t - a - j_t \\ &\geq 2t + a + j_t + 1 + 2r_{t+1} - 2t - a - j_t \\ &= 2r_{t+1} + 1 \geq 2r_{t+1} + 1 - a,\end{aligned}$$

where the third inequality holds by induction and in the second inequality we used the fact that B_{t+1} covers T_{j_t} completely, whereas B_t did not, hence there is at least one vertex in T_{j_t} uncovered by B_t but covered by B_{t+1}. We therefore obtain a contradiction to the assumption that no B_t covers nicely more than $2r_t - a$ vertices. Hence $j_{t+1} \leq 2r_{t+1} + 2 - 2(t+1) - a$ and the first part holds.

To prove the second part we shall consider two cases depending on j_{t+1}.

It is easy to check then when $j_{t+1} = j_t$ then B_{t+1} covers at least $2(r_{t+1} - r_t)$ more vertices of T_{j_t} than B_t. Therefore

$$|T_{j_{t+1}} \cap B_{t+1}| \geq |T_{j_{t+1}} \cap B_t| + 2(r_{t+1} - r_t) \geq |T_{j_t} \cap B_t| + 2.$$

Hence, again by the induction hypothesis,

$$\begin{aligned}\sum_{i=1}^{j_{t+1}} |T_i \cap B_{t+1}| &\geq \sum_{i=1}^{j_t} |T_i \cap B_t| + 2 \geq 2t + a + j_t + 2 \\ &= 2(t+1) + a + j_{t+1}.\end{aligned}$$

On the other hand, when $j_t < j_{t+1}$ then $T_{j_t} \cap B_{t+1} = T_{j_t}$ and hence $|T_{j_t} \cap B_{t+1}| \geq |T_{j_t} \cap B_t| + 1$. Observe that $|T_{j_{t+1}} \cap B_{t+1}| \geq 2r_{t+1} + 2 - j_{t+1} \geq 2r_{t+1} + 2 - (2r_{t+1} + 2 - 2(t+1) - a) = 2(t+1) + a \geq 2,$

by the first part (which we already proved for $t+1$), and hence

$$\begin{aligned}\sum_{i=1}^{j_{t+1}} |T_i \cap B_{t+1}| &= \sum_{i=1}^{j_t} |T_i \cap B_{t+1}| + \sum_{i=j_t+1}^{j_{t+1}-1} |T_i \cap B_{t+1}| + |T_{j_{t+1}} \cap B_{t+1}| \\ &\geq \left(\sum_{i=1}^{j_t} |T_i \cap B_t| + 1 \right) + (j_{t+1} - j_t - 1) + 2 \\ &\geq 2t + a + j_t + 2 + j_{t+1} - j_t = 2(t+1) + a + j_{t+1}.\end{aligned}$$

Which completes the proof. □

We have an immediate corollary to that lemma which, roughly speaking, says that given a tree and a "big" set of distinct non-negative integers there is a ball of radius in that set which covers nicely "many" vertices of the tree.

Corollary 5.10. Let T be a tree and $X = \{r_0, \cdots, r_s\}$ be a set of non-negative integers with $r_i < r_{i+1}$. Then either

1. T can be covered by a ball of radius r_s, or
2. For $a = \max\{r_s + 1 - 2s, 0\}$ there is a ball B of radius $r \in X$ which covers nicely at least $2r + 1 - a$ vertices.

Proof. Suppose for contradiction that T cannot be covered by a ball of radius r_s and for $a = \max\{r_s + 1 - 2s, 0\}$ there is no set of $2r + 1 - a$ vertices which can be covered nicely by a ball of radius $r \in X$. It follows that $r_s < \frac{diam(T)-1}{2}$, hence we can apply Lemma 5.9 and conclude that $r_t \geq 2t + a$, for every $t \in \{0, \cdots, s\}$. In particular $r_s \geq 2s + a \geq 2s + r_s + 1 - 2s = r_s + 1$, which is a contradiction. □

Lemma 5.11. Fix M. Let T be any tree and X be a set of distinct non-negative integers with $\max X \leq M$. Suppose T cannot be covered by balls with radii from X. Then, using balls of radii in X, we can cover at least

$$\sum_{r \in X} (2r+1) - \frac{M^2}{4}$$

vertices of T.

Proof. We shall first show that in the case when the cardinality of X is not too large, i.e., when $|X| \leq \frac{M+1}{2}$, we can cover nicely at least

$$\sum_{r \in X}(2r+1)+|X|^2-|X|M$$

vertices of T. The proof will use induction on the cardinality of $|X|$. The base case $|X|=0$ is trivial. Suppose the result holds for any set of non-negative integers X' with $\max X' \leq M$ and $|X'|<|X|$. It follows from the assumption that $|X| \leq \frac{M+1}{2}$ and Corollary 5.10, that there is a ball B of radius $r \in X$ which, for $a \leq M+1-2|X|$, covers nicely at least $2r+1-a \geq 2r+2|X|-M$ vertices of T. Let A be a set of $2r+2|X|-M$ vertices which are covered nicely by B and let $T' = T \setminus A$ and $X' = X \setminus \{r\}$. By the induction hypothesis we can cover nicely $\sum_{r' \in X'}(2r'+1)+|X'|^2-|X'|M$ vertices in T' using balls of radii from X'. Hence in total we can cover at least

$$\begin{aligned} &\sum_{r' \in X'}(2r'+1)+|X'|^2-|X'|M+2r+2|X|-M \\ =&\sum_{r' \in X'}(2r'+1)+|X|^2+2r+1-|X|M \\ =&\sum_{r \in X}(2r+1)+|X|^2-|X|M \end{aligned}$$

vertices of T with balls of radii in X. This finishes the claim. An instant corollary of that claim is that when $|X| \leq \frac{M+1}{2}$ then the conclusion of the lemma holds, as

$$|X|^2-|X|M=|X|\left(|X|-M\right) \geq -\frac{M^2}{4}.$$

Suppose now that $|X| \geq \frac{M+1}{2}$. Observe that in this case it follows from Corollary 5.10 that there is a ball of radius $r \in X$ which covers nicely at least $2r+1$ vertices of T. Applying this observation iteratively we can conclude that there is a subset

Y of X such that $|X \setminus Y| \leq \frac{M+1}{2}$ and we can cover nicely $L' \geq \sum_{r\in Y}(2r+1)$ vertices of T. Let T' be the uncovered subtree of T (hence $|T'| \leq T - L'$) and $X' = X \setminus Y$. Since $|X'| \leq \frac{M+1}{2}$ we can use the claim made in the paragraph above to conclude that we can cover nicely $L'' \geq \sum_{r\in X'}(2r+1) + |X'|^2 + |X'|M$ vertices of T'. In total we have covered at least

$$\begin{aligned} L' + L'' &\geq \sum_{r\in Y}(2r+1) + \sum_{r\in X\setminus Y}(2r+1) + |X'|^2 - |X'|M \\ &= \sum_{r\in X}(2r+1) + |X'|^2 - |X'|M \\ &\geq \sum_{r\in X}(2r+1) - \frac{M^2}{4}. \end{aligned}$$

□

We can now easily deduce Theorem 5.2 from Lemma 5.11.

Proof of Theorem 5.2. Let $X = \{0, \ldots, m-1\}$, where $m = \left\lceil \sqrt{\frac{4}{3}n} \right\rceil$. Suppose for contradiction that T cannot be burned in m rounds. Then, by Lemma 5.11, we can cover $\sum_{r=0}^{m-1}(2r+1) - \frac{(m-1)^2}{4} \geq m^2 - \frac{m^2}{4} = \frac{3}{4}m^2 \geq n$ vertices using the balls of radii at most m. This gives a contradiction. □

CHAPTER 6

ON POSSIBLE NUMBERS OF COPIES OF A FIXED GRAPH

In this chapter we investigate the set T_n of possible number of triangles in a graph on n vertices. The first main result says that every natural number less than $\binom{n}{3} - \left(\sqrt{2}+o(1)\right)n^{3/2}$ belongs to T_n. On the other hand, we show that there is a number $m = \binom{n}{3} - \left(\sqrt{2}+o(1)\right)n^{3/2}$ which is not a member of T_n. In addition, we prove that there are two interlacing sequences $\binom{n}{3} - \left(\sqrt{2}+o(1)\right)n^{3/2} = c_1 \le d_1 \le c_2 \le d_2 \le \cdots \le c_s \le d_s = \binom{n}{3}$ with $|c_t - d_t| = n-2-\binom{s-t+1}{2}$ such that $(c_t, d_t) \cap T_n = \emptyset$ for all t. Moreover, we obtain a generalization of these results for the set of possible number of copies of a connected graph H in a graph on n vertices. This work is joint with Teeradej Kittipassorn.

6.1 Introduction

We ask the following natural question: given a graph H and a natural number n, what are the possible values of m such that there exists a graph on n vertices with *exactly* m copies of H? Surprisingly, very little is known about this problem.

A question of this flavor was first considered by Kittipassorn and Mészáros [56] who studied the set F_n of possible number of *frustrated triangles*, i.e. triples of vertices inducing an odd number of edges. They proved that about two thirds of the numbers in $\left[0, n^{3/2}\right]$ do not appear in F_n and every even number between $(1+o(1))n^{3/2}$ and $\binom{n}{3} - (1+o(1))n^{3/2}$ is a member of F_n for sufficiently large and even n.

Much more attention has been given to the problem of maximizing or minimizing the number of subgraphs of certain type in graphs of given number of vertices and edges. For example, Rademacher proved that every graph with $\lfloor n^2/4 \rfloor + 1$ edges contains at least $\lfloor n/2 \rfloor$ triangles. Erdős [29] posed a conjecture, which was later proved by Lovasz and Simonovits [64], that a graph of size $\lfloor n^2/4 \rfloor + k$ contains at least $k\lfloor n/2 \rfloor$ triangles if

$k < n/2$. On the other hand, Alon [2] investigated the maximum number of subgraphs isomorphic to some given graph where the maximum is taken over all graphs of certain size. We refer interested readers to [42, 27, 62, 31, 28, 30, 70, 71] for similar results.

We consider a fixed connected graph H. For a graph G, we define $k_H(G)$ to be the number of copies of H in G. The main object of interest of this chapter is

$$S_H^{(n)} = \{k_H(G) : |G| = n\},$$

the set of possible number of copies of H in a graph on n vertices. Our first result says that almost every number (in the appropriate range) is realizable as a number of copies of H in some graph of order n.

Theorem 6.1. As $n \to \infty$, the following holds

$$[0, (1 - o(1))k_H(K_n)] \subset S_H^{(n)}.$$

It is not unreasonable to expect that $o(1)$ above could be replaced by 0. As it happens, this is not the case. Before we state the second result, let us introduce some notations which play a major role in the rest of the chapter. For a graph $G = (V,E)$ of order n, let $f_H(G)$ be the number of subgraphs H of $K(V)$ isomorphic to F such that $E(H) \cap E \neq \emptyset$, where $K(V)$ denotes the complete graph on the vertex set V. For instance, $f_{K_3}(G)$ is the number of triples of vertices in G inducing at least one edge. Observe that $k_H(G) = k_H(K_n) - f_H(\overline{G})$, and therefore we shall work instead with the complement of G which is easier to draw when G is dense. We shall write $a_H^{(n)}(t) = f_H(S_t^{(n)})$ and $b_H^{(n)}(t) = f_H(M_t^{(n)})$ where $S_t^{(n)}$ is a graph on n vertices with t edges forming a star with $t \leq n-1$ and $M_t^{(n)}$ is a graph on n vertices with t edges forming a matching with $t \leq \lfloor n/2 \rfloor$. We prove the existence of some forbidden intervals in $S_H^{(n)}$.

Theorem 6.2. For all sufficiently large n and all $t \geq 0$, we have

$$\left| \left(k_H(K_n) - a_H^{(n)}(t+1), k_H(K_n) - b_H^{(n)}(t) \right) \cap S_H^{(n)} \right| = o(n^{h-2}).$$

We shall see in Section 6.4 that when $t < c\sqrt{n}$, where c is some nonnegative constant depending only on H, then $a_H^{(n)}(t+1) - b_H^{(n)}(t)$ is of order n^{h-2}.

In the case when H is a triangle, we prove a sharp analog of Theorems 6.1 and 6.2.

Theorem 6.3. The following hold.

(i) $\left[0, \binom{n}{3} - (\sqrt{2}+o(1))n^{3/2}\right] \subset S_{K_3}^{(n)}$ as $n \to \infty$.

(ii) $\left(\binom{n}{3} - a_{K_3}^{(n)}(t+1), \binom{n}{3} - b_{K_3}^{(n)}(t) \right) \cap S_{K_3}^{(n)} = \emptyset$ for all $n, t \geq 0$.

We shall remark that $a_{K_3}^{(n)}(t) = t(n-2) - \binom{t}{2}$ and $b_{K_3}^{(n)}(t) = t(n-2)$, and so it is easy to check that the interval in the second part of the theorem is not empty as long as $t \lesssim \sqrt{2n}$, whence there exists a number $m = \binom{n}{3} - \left(\sqrt{2}+o(1)\right) n^{3/2}$ which is not a member of T_n. Therefore the first part of Theorem 6.3 is sharp.

The rest of this chapter is organized as follows. In Section 6.2, we prove some preliminary lemmas. In Section 6.3, we prove the sharp results, Theorem 6.3 for triangles. The proofs of Theorems 6.1 and 6.2 are presented in Section 6.4. We conclude the chapter in Section 6.5 with some open problems.

6.2 Complete graphs

In this section we shall consider the case when $H = K_r$ is a complete graph and prove some basic lemmas which we shall later use to prove some of our main theorems.

Let $P_{n,r}$ be the set of possible number of copies of K_r in an r-partite graph on n vertices. Clearly $P_{n,r} \subseteq S_{K_r}^{(n)}$.

We shall start by showing that the first $\left(\left\lfloor \frac{n-r^2}{r} \right\rfloor \right)^r$ natural numbers are realizable, i.e. for every $k \leq \left(\left\lfloor \frac{n-r^2}{r} \right\rfloor \right)^r$ there exists an r-partite graph on n vertices with exactly k copies of K_r.

Lemma 6.4. For natural numbers $n \geq r \geq 2$ and any non-negative integer $k \leq \left(\left\lfloor \frac{n-r^2}{r} \right\rfloor\right)^r$ there is an r-partite graph with k copies of K_r.

Proof. We shall use induction on r. The base case $r = 2$ is trivial. Suppose the assertion holds for some $r \geq 2$. Let $G = (V_1 \sqcup V_2 \sqcup V_3, E)$ where $|V_1| = \lceil \frac{r}{r+1}(n-r) \rceil$, $|V_2| = \lfloor \frac{1}{r+1}(n-r) \rfloor$, every vertex in V_1 is joint to every vertex in V_2, and V_3 induces K_r. By the induction hypothesis we can replace V_1 by an r-partite graph having k copies of K_r, therefore obtaining an $(r+1)$-partite graph having $\lfloor \frac{1}{r+1}(n-r) \rfloor k$ copies of K_{r+1}, for any $k \leq \left(\frac{\frac{r(n-r)}{r+1} - r^2}{r} \right)^r = \left(\frac{n-r}{r+1} - r \right)^r$. Therefore, we get an increasing sequence in $P_{n,r+1}$ starting with 0 and ending with
$\lfloor \frac{n-r}{r+1} \rfloor \left(\lfloor \frac{n-r}{r+1} - r \rfloor \right)^r \geq \left(\lfloor \frac{n-r}{r+1} - r \rfloor \right)^{r+1} = \left(\left\lfloor \frac{n-r(r+2)}{r+1} \right\rfloor \right)^{r+1} \geq \left(\left\lfloor \frac{n-(r+1)^2}{r+1} \right\rfloor \right)^{r+1}$ such that the difference between consecutive terms is equal to $|V_2|$. To obtain the missing numbers between consecutive terms, notice that it is enough to join needed number of vertices in V_2 to every vertex in V_3. Clearly all graphs in the sequence are $(r+1)$-partite. □

We shall fix $n, r \geq 2$ and for brevity write $f(G) = f_{K_r}(G)$, $a_t = a^{(n)}_{K_r}(t)$, $b_t = b^{(n)}_{K_r}$. Recall that the number of copies of K_r in G is equal to $\binom{n}{r} - f(\overline{G})$ where $f(G)$ is the number of r-sets of vertices in G which induce at least one edge. Therefore, we shall later work instead with the complement of G which is easier to deal with when G is dense. Observe that $a_t = \sum_{i=1}^{t} \binom{n-1-i}{r-2}$ and hence $a_{t+1} - a_t = \binom{n-1-(t+1)}{r-2}$.

Lemma 6.5. For a graph G on n vertices and $e \leq \frac{n-1}{2}$ edges, $f(G) \in [a_e, b_e]$.

Proof. We shall show this by induction on the number of edges e. In the base case when $e \leq 1$, there is nothing to show. For $e > 1$, assume that $f(G') \in [a_{e-1}, b_{e-1}]$ for any graph G' on n vertices and $e-1$ edges.

First we shall show that $f(G) \leq b_e$. Take an edge $xy \in G$ such that $d(y) > 1$ and an isolated vertex $w \in G$. Let G' be a graph obtained by removing xy from G and replacing it by xw. Notice that $f(G') \geq f(G)$, hence repeating this process for any nonindependent edge, we eventually obtain a matching, without decreasing the value of f.

To show that $f(G) \geq a_e$, pick an edge $xy \in G$ and let G' be a graph obtained by removing xy from G. We shall show that $f(G) - f(G') \geq \binom{n-1-e}{r-2} = a_e - a_{e-1}$ which will complete the proof, as then $f(G) \geq f(G') + a_e - a_{e-1} \geq a_e$, by the induction hypothesis applied to G'. Let $A = V(G) \setminus (N_G(x) \cup N_G(y))$ (observe that $x, y \neq A$). Write M for a largest independent set contained in A and e_A for the number of edges induced by A. It is easy to show that $|M| \geq |A| - e_A$. Therefore
$f(G) - f(G') \geq \binom{|M|}{r-2} \geq \binom{n-|N(x) \cup N(y)| - e_A}{r-2} \geq \binom{n-e-1}{r-2}$. □

We remark that Lemma 6.5 does not imply that $f(G) \notin (b_t, a_{t+1})$. However, the result follow immediately from the monotonicity of f.

Lemma 6.6. For any graph G on n vertices and any $t \geq 0$, $f(G) \notin (b_t, a_{t+1})$.

Proof. We shall write $t_{max} = \max\{t : b_{t-1} + 1 < a_t\}$ for the last t where there is a gap between the intervals $[a_{t-1}, b_{t-1}]$ and $[a_t, b_t]$. It is enough to show that $f(G) \in [a_t, b_t]$ for some $t \leq t_{max}$ or $f(G) \geq a_{t_{max}}$. By Lemma 6.5, we are done if $e(G) \leq t_{max}$. So we can assume that $e(G) > t_{max}$. Let G' be a graph obtained from G by deleting some edges until there are exactly t_{max} edges left. By monotonicity of f, we have $f(G) > f(G') \geq a_{t_{max}}$ by Lemma 6.5. □

We remark that $t_{max} = \Theta(\sqrt{n})$.

6.3 Triangle

We shall now consider the case when $H = K_3$, i.e. when H is a triangle. For brevity, let us write $T_n = S^{(n)}_{K_3}$, $f(G) = f_{K_3}(G)$, $a_t = a^{(n)}_{K_3}(t)$ and $b_t = b^{(n)}_{K_3}(t)$.

In order to improve Lemma 6.4 to Theorem 6.3(i), let us change the direction to Theorem 6.3(ii) and look for nonmembers of T_n. Recall that the number of triangles in G is equal to $\binom{n}{3} - f(\overline{G})$ where $f(G)$ is the number of triples of vertices in G which induce at least one edge. Therefore, we shall work instead with the complement of G which is easier to deal with when G is dense. Notice that we have a simple formula $f(G) = e(G)(n-2) - n_c + n_t$ where n_c is the number of cherries (i.e. paths with two edges, P_2) and n_t is the number of triangles in G. This comes from the fact that each edge is contained in exactly $n-2$ triples, but we double count the triples which contain more than one edge. Using this formula it is easy to see that $a_t = t(n-2) - \binom{t}{2}$ and $b_t = t(n-2)$.

We have shown in Lemma 6.6 that $f(G) \notin (b_t, a_{t+1})$ for all $t \geq 0$. On the other hand, we shall prove that any number bigger than $\left(\sqrt{2}+o(1)\right)n^{3/2}$ is realizable.

Lemma 6.7. For any sufficiently large n, if $m \in \left[\left(\sqrt{2}+o(1)\right)n^{3/2}, (1-o(1))\binom{n}{3}\right]$ then there is a graph G on n vertices such that $f(G) = m$.

Proof. Given natural number n and an integer $m \in \left[\left(\sqrt{2}+o(1)\right)n^{3/2}, (1-o(1))\binom{n}{3}\right]$ we shall construct graph G on n vertices such that $f(G) = m$. Let us partition G into four parts, so $V(G) = V_1 \sqcup V_2 \sqcup V_3 \sqcup V_4$ where

- $G[V_1]$ is empty of order $n' = n - 2\sqrt{2n} - 4(8n)^{1/4}$,
- $G[V_2], G[V_3], G[V_4]$ are matchings of sizes $\sqrt{2n}$, $(8n)^{1/4}$, $(8n)^{1/4}$ respectively,
- there are no edges between the classes, i.e. $E(V_i, V_j) = \emptyset$, for $i \neq j$.

We shall consider a sequence of graphs obtained by adding edges one by one to $G[V_1]$, i.e. $G_0 = G$ and $G_i[V_1] = G_{i-1}[V_1] \cup e$ for some edge $e \notin G_{i-1}[V_1]$, for any $0 < i \leq \binom{n'}{2}$. Observe that for sufficiently large n we have $f(G_0) = (n-2)\left(\sqrt{2n} - 2(8n)^{1/4}\right) = \left(\sqrt{2}+o(1)\right)n^{3/2}$ and $f(G_{\binom{n'}{2}}) = (1+o(1))\binom{n}{3}$. Therefore, in order to prove the lemma, it suffices to show that any number in the interval $(f(G_{i-1}), f(G_i))$ is realizable. We shall achieve that by moving edges within each of $G_i[V_2], G_i[V_3], G_i[V_4]$, but not across them. As an edge is contained in $n-2$ triples, it follows easily that $f(G_i) - f(G_{i-1}) \leq n-2$.

Let us fix $i \geq 1$. By construction, $V(G_i) = V_1 \sqcup V_2 \sqcup V_3 \sqcup V_4$. Let $\left\{x_1y_1, \cdots x_{\sqrt{2n}}y_{\sqrt{2n}}\right\}$ be the matching inside V_2. We shall construct another sequence $G_{i,1}, G_{i,2}, \cdots, G_{i,\sqrt{2n}}$ where $G_{i,k+1}$ is obtained from $G_{i,k}$ by deleting edge $x_{k+1}y_{k+1}$ and adding edge x_1y_{k+1} (see Figure 1). Notice that

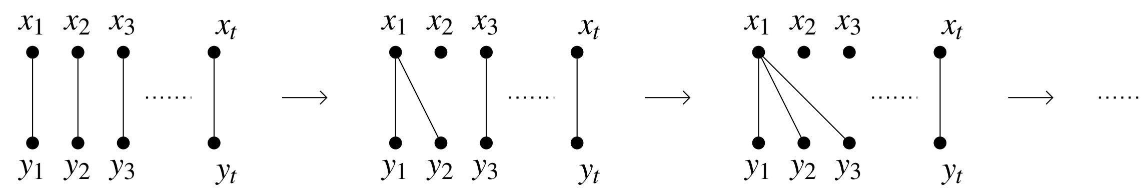

Figure 1: Star accumulation of V_2 and V_3.

$f(G_{i,\sqrt{2n}}) = f(G_i) - \binom{\sqrt{2n}}{2} = f(G_i) - \frac{\sqrt{2n}(\sqrt{2n}-1)}{2} = f(G_i) - n + \sqrt{2n}/2$ and $f(G_{i,k}) - f(G_{i,k+1}) = k$. Hence we obtain a refinement of the sequence with the gaps between consecutive terms bounded by $\sqrt{2n}$.

Let us fix $i \geq 1$ and $j \geq 2$. Let $\left\{x_1y_1, \cdots x_{(8n)^{1/4}}y_{(8n)^{1/4}}\right\}$ be the matching inside V_3. We shall construct another sequence $G_{i,j,1}, G_{i,j,2}, \cdots, G_{i,j,(8n)^{1/4}}$ where $G_{i,j,k+1}$ is obtained from $G_{i,j,k}$ by deleting edge $x_{k+1}y_{k+1}$ and adding edge x_1y_{k+1}. Notice that $f(G_{i,j,(8n)^{1/4}}) = f(G_{i,j}) - \binom{(8n)^{1/4}}{2} = f(G_i) - \frac{(8n)^{1/4}((8n)^{1/4}-1)}{2} = \sqrt{2n} - (8n)^{1/4}/2$ and $f(G_{i,j,k}) - f(G_{i,j,k+1}) = k$. Hence we obtain a refinement of the sequence with the gaps between consecutive terms bounded by $(8n)^{1/4}$.

Finally, let us fix i, j, k. Let $\left\{x_1y_1, \cdots x_{(8n)^{1/4}}y_{(8n)^{1/4}}\right\}$ be the matching inside V_4. We shall construct another sequence $G_{i,j,k,1}, G_{i,j,k,2}, \cdots, G_{i,j,k,(8n)^{1/4}}$ where $G_{i,j,k,l+1}$ is obtained from $G_{i,j,k,l}$ by deleting edge $x_{k+1}y_{k+1}$ and adding edge y_ky_{k+1} (see Figure 2).

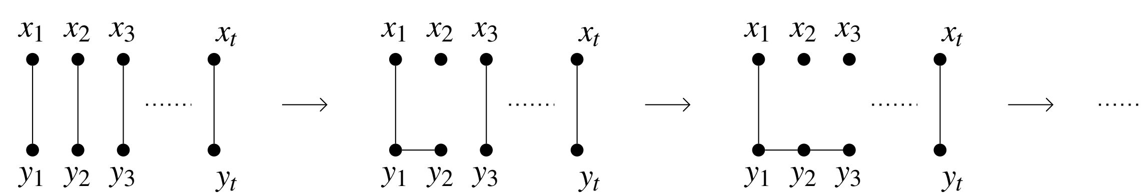

Figure 2: Path accumulation of V_4.

Notice that $f(G_{i,j,k,(8n)^{1/4}}) = f(G_{i,j,k}) - (8n)^{1/4}$ and $f(G_{i,j,k,l}) - f(G_{i,j,k,l+1}) = 1$. Hence we obtain a refinement of the sequence with no gaps.

□

We are now ready to deduce Theorem 6.3 from Lemmas 6.4, 6.6 and 6.7.

Proof of Theorem 6.3. (i) Recall that the number of triangles in G is equal to $\binom{n}{3} - f(\overline{G})$. Therefore, Lemma 6.7 implies that $\left[o(\binom{n}{3}), \binom{n}{3} - (\sqrt{2}+o(1))n^{3/2}\right] \subset T_n$. Together with Lemma 6.4 which, for $r=3$, says that $\left[0, \left(\frac{n-9}{3}\right)^3\right] \subset T_n$, we conclude that $\left[0, \binom{n}{3} - \left(\sqrt{2}-o(1)\right)n^{3/2}\right] \subseteq T_n$ for sufficiently large n.

(ii) Since the number of triangles in G is equal to $\binom{n}{3} - f(\overline{G})$, we obtain, using Lemma 6.6, that $\left(\binom{n}{3} - a_{t+1}, \binom{n}{3} - b_t\right) \cap T_n = \emptyset$ for all $n, t \geq 0$. □

6.4 General H

Now, let us consider the case when H is an arbitrary connected graph on $h \geq 3$ vertices. We shall start by showing that when n goes to infinity then the first $(1-o(1))k_H(K_n)$ numbers are realizable.

Our strategy will be to recursively partition the vertex set into two subsets and modify the edges between vertices in each of the classes, but without adding diagonal edges. Let $g_H = g_H^{(n)}$ be the maximum number of new copies of H obtained by adding an edge to a graph, over all graphs on n vertices. We claim that there is a constant $c_H > 0$ such that $g_H \leq c_H n^{h-2}$. Indeed, a new copy must contain both the endvertices of the newly added edge, and there are $\binom{n-2}{h-2}$ h-sets of vertices in G containing two fixed vertices, and each h-set may contain at most c'_H copies of H, for some c'_H independent of n, therefore $g_H \leq c'_H\binom{n-2}{h-2} \leq c_H n^{h-2}$.

The next two lemmas are needed in our construction.

Lemma 6.8. If $[0, cn^{\alpha}] \subset S_H^{(n)}$ for all sufficiently large n, where $\alpha \leq h-2$, then for all sufficiently large n and some new constant $c_1 > 0$,

$$\left[0, c_1 n^{\alpha h/(h-2)}\right] \subset S_H^{(n)}.$$

Proof. Consider an empty graph G with vertex set $V = V_1 \sqcup V_2$, where $n_1 = |V_1| = c' n^{\alpha/(h-2)}$ and $n_2 = |V_2| = n - |V_1|$, where $c' > 0$ will be chosen later. Let $G_0 = G$ and let G_{i+1} be a graph obtained by adding an edge between vertices of V_1 in G_i, then

$$k_H(G_{i+1}) - k_H(G_i) \leq g_H(n_1) \leq c_H n_1^{h-2} = c_H c'^{h-2} n^{\alpha}.$$

Therefore we obtain an increasing sequence in $S_H^{(n)}$ starting with 0 and ending with $k_H(G_{\binom{n_1}{2}})$ such that the differences between consecutive terms are at most $c_H c'^{h-2} n^{\alpha}$. We shall modify $G_i[V_2]$ to obtain the missing numbers between consecutive terms. By the hypothesis, we can modify $G_i[V_2]$, to obtain $G_i'[V_2]$ containing any number k of copies of H, where $k \in \left[0, c n_2^{\alpha}\right]$. Hence it suffices to find $c' > 0$ such that $c_H c'^{h-2} n^{\alpha} < c n_2^{\alpha}$. Let us consider two cases depending on α.

1. if $\alpha = h-2$ then $c n_2^{\alpha} = c\left(n - c'n\right)^{\alpha} = c(1-c')^{\alpha} n^{\alpha}$, therefore it suffices to choose $c' > 0$ such that $c_H c'^{h-2} < c(1-c')$. Hence we have $g_H^{(n_1)} < c n_2^{\alpha}$.

2. if $\alpha < h-2$ then $c n_2^{\alpha} = c\left(n - o(n)\right)^{\alpha} \sim c n^{\alpha}$, hence if we choose $c' > 0$ such that $c' < \left(c/c_H\right)^{1/(h-2)}$, then for sufficiently large n we will have $g_H^{(n_1)} < c n_2^{\alpha}$.

Therefore, any number less than $k_H(G_{\binom{n_1}{2}}) = k_H(K_{n_1}) > c'' n_1^h > c_1 n^h$ is realizable. □

From the next lemma we learn that for sufficiently large n we can construct a graph on n vertices with k of copies of H for any $k \leq (1-o(1)) k_H(K_n)$

Lemma 6.9. If $[0, c n^{\alpha}] \subset S_H^{(n)}$ for sufficiently large n where $\alpha > h-2$ then for sufficiently large n,

$$[0, (1-o(1)) k_H(K_n)] \subset S_H^{(n)}.$$

Proof. We shall proceed similarly as in Lemma 6.8. Choose $\beta \in ((h-2)/\alpha, 1))$ and let $n_2 = |V_2| = n^\beta$ and $n_1 = |V_1| = n - |V_2|$. Notice that $g_H^{(n_1)} = O(n^{h-2})$ and by the hypothesis we can modify $G[V_2]$ to obtain any number of copies of H up to n_2^α, where $n_2^\alpha = \omega(n^{h-2})$. Therefore any number in the interval $[0, k_H(n_1)]$ is realizable. But $n_1 = (1-o(1))n$, whence $k_H(n_1) = (1-o(1))k_H$. □

We shall use Lemmas 6.8 and 6.9 to prove one of the two main theorems of this section.

Proof of Theorem 6.1. We start by showing that trivially $[0, \lfloor n/h \rfloor] \subset S_H^{(n)}$. To achieve that notice that for any $k \leq \lfloor n/h \rfloor$ we can simply construct a graph on n vertices consisting of k disjoint copies of H.

Let k_{max} be the largest integer k such that $(\frac{h}{h-2})^k \leq h$ (note that $(\frac{h}{h-2})^{k_{max}} \in (h-2, h])$. We claim that $\left[0, c_k n^{(h/(h-2))^k}\right] \subset S_H^{(n)}$ for every positive integer $k \leq k_{max}$. We shall show the claim by induction on k. For $k = 0$, we already showed that $[0, c_0 n] \subset S_H^{(n)}$. Suppose, that $\left[0, c_k n^{(h/(h-2))^k}\right] \in S_H^{(n)}$ and $k < k_{max}$. Observe first that $(h/(h-2))^k \leq h-2$, as otherwise $(h/(h-2))^{k_{max}}$ would be greater than h, hence we can apply Lemma 6.8 and conclude that $[0, c_{k+1} n^{(h/(h-2))^{k+1}}] \in S_H^{(n)}$ for large enough n. Note that we apply Lemma 6.8 only finitely many times hence n remains finite.

Therefore for n large enough we have $[0, cn^\alpha] \subset S_H^{(n)}$ with $\alpha = (\frac{h}{h-2})^{k_{max}} \in (h, h-2]$, hence we can apply Lemma 6.9 and conclude that for n sufficiently large $[0, (1-o(1))k_H] \subset S_H^{(n)}$.

□

Let us recall few major notations. For a graph $G = (V, E)$ of order n, let $f_H(G)$ be the number of subgraphs H of $K(V)$ isomorphic to F such that $E(H) \cap E \neq \emptyset$, where $K(V)$ denotes the complete graph on the vertex set V. Then the number of copies of H in G is equal to $k_H(K_n) - f_H(\overline{G})$.

In the next lemma we shall describe the formula for f_H.

Lemma 6.10. For any graph H on h vertices and G on n vertices, we have

$$f_H(G) = c_H e(G)\binom{n-2}{h-2} - \sum_{k=2}^{e(H)} \sum_{\substack{e(F)=k \\ \delta(F)>0}} (-1)^{k+1} c_H(F) k_F(G) \binom{n-|V(F)|}{h-|F|}.$$

Proof. Let $E(G) = \{e_1, \cdots, e_m\}$, where $m = e(G)$. We define $A_i = \{F : e_i \in E(F), F \cong H\}$ to be the set of subgraphs of the complete graph isomorphic to H containing the edge e_i. Notice that $f_H(G) = |\bigcup_{i=1}^m A_i|$, therefore by the inclusion-exclusion principle we can write

$$f_H(G) = \sum_{k=1}^{e(G)} \sum_{i_1<\ldots<i_k} (-1)^{k+1} |A_{i_1} \cap \ldots \cap A_{i_k}|.$$

For a graph F on at most h vertices, let $c_H(F)$ be the number of copies of H in the complete graph K_h containing all the edges of a fixed subgraph of the complete graph K_h, isomorphic to F. Let $F = G[e_{i_1}, \ldots, e_{i_k}]$ be the graph induced by the edges $e_{i_1}, \cdots, e_{i_k}$. Observe that $|A_{i_1} \cap \cdots \cap A_{i_k}| = c_H(F)\binom{n-|F|}{h-|F|}$, therefore

$$\begin{aligned}
f_H(G) &= \sum_{k=1}^{e(G)} \sum_{i_1<\ldots<i_k} (-1)^{k+1} |A_{i_1} \cap \ldots \cap A_{i_k}| \\
&= \sum_{k=1}^{e(G)} \sum_{\substack{F\subseteq H \\ e(F)=k \\ \delta(F)>0}} (-1)^{k+1} k_F(G) c_H(F) \binom{n-|F|}{h-|F|} \\
&= \sum_{k=1}^{e(H)} \sum_{\substack{F\subseteq H \\ e(F)=k \\ \delta(F)>0}} (-1)^{k+1} k_F(G) c_H(F) \binom{n-|F|}{h-|F|} \\
&= c_H e(G)\binom{n-2}{h-2} - \sum_{k=2}^{e(H)} \sum_{\substack{e(F)=k \\ \delta F>0}} (-1)^{k+1} c_H(F) k_F(G) \binom{n-|F|}{h-|F|}.
\end{aligned}$$

□

The following easy lemma gives us an upper bound for $k_F(G)$.

Lemma 6.11. If H is a graph on h vertices with no isolated vertices then for every graph G on e edges the number of copies of H in G is at most e^{h-1}.

Proof. We shall proceed by induction. The base case $h = 2$ is trivial. Assume then that $h > 2$ and let us consider two cases. If H is a matching on $h = 2l$ vertices then the result follows easily - the number of copies of H in G is at most $\binom{e}{l} \le e^l \le e^{2l-1} = e^{h-1}$. In the other case, when H is not a matching, there exists a vertex $v \in H$ such that $H' = H - v$ has no isolated vertices. By the induction hypothesis there are at most $e^{h'-1}$ copies of H' in G, where $h' = |H'| = h-1$. Each copy of H' in G can be extended to at most e copies of H in G, since by assumption v must be adjacent to some vertex of H'. Therefore there are at most $e^{h'} = e^{h-1}$ copies of H in G. □

Lemma 6.12. We have

$$f_H(G) = c_H e(G)\binom{n-2}{h-2} - c_H(P_2)k_{P_2}(G)\binom{n-3}{h-3} + c_H(K_3)k_{K_3}(G)\binom{n-3}{h-3} + o(n^{h-2}).$$

Proof. Let us consider the term $c_H(F)k_F(G)\binom{n-|F|}{h-|F|}$ where F is a graph on at least four vertices. It follows from Lemma 6.11 that $k_F(G) \le e^{|F|-1}$ and therefore $k_F(G)\binom{n-|F|}{h-|F|} \le e^{|F|-1}n^{h-|F|}$. Hence, under the assumption that $e = O(n^{1/2})$, we have $k_F(G)\binom{n-|F|}{h-|F|} = O\left(n^{1/2|F|-1/2}n^{h-|F|}\right) = O\left(n^{h-1/2-1/2|F|}\right) = O\left(n^{h-5/2}\right) = o(n^{h-2})$. As there are only three graphs on fewer than four vertices with no isolated vertices, namely K_2, K_3 and P_2, and the number of terms in the summation in $f(G)$ depends only on H, we can write

$$f_H(G) = c_H e(G)\binom{n-2}{h-2} - c_H(P_2)k_{P_2}(G)\binom{n-3}{h-3} + c_H(K_3)k_{K_3}(G)\binom{n-3}{h-3} + o(n^{h-2}).$$

□

The next lemma, which we shall later use to prove that there are gaps in $S_H^{(n)}$, tells us that for sufficiently large n, stars and matchings are asymptotically extremal examples of graphs for $f_H(G)$, i.e. for a graph G on t edges we have

$$a_H^{(n)}(t)-o(n^{h-2})=f_H(S_t^{(n)})-o(n^{h-2})\leq f_H(G)\leq f_H(M_t^{(n)})+o(n^{h-2})=b_H^{(n)}(t)+o(n^{h-2}).$$

Lemma 6.13. Let G be a graph on n vertices with $e=O(n^{1/2})$ edges. Then the following hold.

1. $f_H(G)\geq a_H^{(n)}(e)-o(n^{h-2})=c_He\binom{n-2}{h-2}-c_H(P_2)\binom{e}{2}\binom{n-3}{h-3}+o(n^{h-2})$; and
2. $f_H(G)\leq b_H^{(n)}(e)+o(n^{h-2})=c_He\binom{n-2}{h-2}+o(n^{h-2})$,

as n goes to infinity.

Proof. This an immediate corollary of Lemma 6.12. Observe that that $k_{K_3}(S_e^{(n)})=0$, as stars contain no triangles, and $k_{P_2}(S_e^{(n)})=\binom{e}{2}$. Therefore, by Lemma 6.12

$$a_H^{(n)}(e)=c_He\cdot\binom{n-2}{h-2}-c_H(P_2)\binom{e}{2}\binom{n-3}{h-3}+o(n^{h-2}).$$

On the other hand, matchings contain no copies of K_3 nor P_2, hence, again, by Lemma 6.12

$$b_H^{(n)}(e)=c_He\cdot\binom{n-2}{h-2}+o(n^{h-2}).$$

Now, since for any graph G on e edges we have $k_{P_2}(G)\leq\binom{e}{2}$ it follows from the above estimate on $a_H^{(n)}(e)$ and from Lemma 6.12 that

$$f_H(G)\geq c_He\cdot\binom{n-2}{h-2}-c_H(P_2)\binom{e}{2}\binom{n-3}{h-3}+o(n^{h-2})\geq a_H^{(n)}(e)-o(n^{h-2}).$$

On the other hand, it follows from easy to see fact that for any graph G we have $c_H(P_2)k_{P_2}(G)\geq c_H(K_3)k_{K_3}(G)$ and Lemma 6.12 that

$$f_H(G)\leq c_He\cdot\binom{n-2}{h-2}+o(n^{h-2})\leq b_H^{(n)}(e)+o(n^{h-2}).$$

□

We can now prove the second main theorem of the section.

Proof of Theorem 6.2. Let $t_{max} = \max\left\{t : b_H^{(n)}(t) < a_H^{(n)}(t+1)\right\}$. We shall first show that $t_{max} = \Theta(\sqrt{n})$. Indeed, we have

$$\begin{aligned} a_H^{(n)}(t+1) - b_H^{(n)}(t) &= c_H \binom{n-2}{h-2} - c_H(P_2)\binom{t+1}{2}\binom{n-3}{h-3} + o(n^{h-2}) \\ &= c_1 n^{h-2} - c_2 t^2 n^{h-3} + o(n^{h-2}), \end{aligned}$$

for some constants $c_1, c_2 > 0$ depending depending only on H. It follows that there is a constant C depending only on H, such that for all sufficiently large n we have $a_H^{(n)}(t+1) - b_H^{(n)}(t) \geq 0$ if $t \geq C\sqrt{n}$ and $a_H^{(n)}(t+1) - b_H^{(n)}(t) \leq 0$ otherwise. From Lemma 6.13 we know that for all sufficiently large n we have

$$f_H(G) \in \left[a_H^{(n)}(t) - o(n^{h-2}), b_H^{(n)}(t) + o(n^{h-2})\right].$$

Therefore, for any $t < t_{max}$, the number of integers m in the interval $\left(b_H^{(n)}(t), a_H^{(n)}(t+1)\right)$ such that there is a graph G on n vertices with $f_H(G) = m$ is at most $o(n^{h-2})$. Whence $\left|(k_H(K_n) - a_{t+1}, k_H(K_n) - b_t) \cap S_H^{(n)}\right| = o(n^{h-2})$, for every t. Notice that when $t < \frac{t_{max}}{2}$, the gap between a_{t+1} and b_t is of the order n^{h-2}, hence

$$\frac{\left|(k_H(K_n) - a_{t+1}, k_H(K_n) - b_t) \cap S_H^{(n)}\right|}{a_{t+1} - b_t} = o(1),$$

as n goes to infinity. □

6.5 Open problems

We conclude this chapter with some open problems that we feel would merit further study. Let $\phi_H^{(n)} = \min\left\{m \geq 0 : m \notin S_H^{(n)}\right\}$ be the smallest nonmember of $S_H^{(n)}$. In this chapter we have proved that $\phi_{K_3}^{(n)} = \binom{n}{3} - \left(\sqrt{2} + o(1)\right) n^{3/2}$. We would like to remark that we were also able to prove the following two results, whose proofs we omit, as they are very similar to the proof of Theorem 6.3.

- $\phi_{P_2}^{(n)} = 3\binom{n}{3} - (4 + o(1))\, n^{3/2}$,
- $\binom{n}{4} - (c + o(1))\, n^{8/3} \leq \phi_{K_4}^{(n)} \leq \binom{n}{4} - \left(\frac{1}{2} + o(1)\right) n^{5/2}$.

The ultimate goal is to determine $S_H^{(n)}$, in particular we ask the following question.

Problem 6.14. What is the asymptotic behavior of $\binom{n}{r} - \phi_{K_r}^{(n)}$?

REFERENCES

[1] *GLPK (GNU Linear Programming Kit).* `https://www.gnu.org/software/glpk/`.

[2] N. ALON, *On the number of subgraphs of prescribed type of graphs with a given number of edges*, Israel Journal of Mathematics, 38 (1981), pp. 116–130.

[3] N. ALON, *Splitting digraphs*, Combin. Probab. Comput., 15 (2006), pp. 933–937.

[4] N. ALON AND K. BERMAN, *Regular hypergraphs, Gordon's lemma, Steinitz' lemma and invariant theory*, Journal of Combinatorial Theory, Series A, 43 (1986), pp. 91–97.

[5] N. ALON, P. PRAŁAT, AND N. WORMALD, *Cleaning regular graphs with brushes*, SIAM Journal on Discrete Mathematics, 23 (2009), pp. 233–250.

[6] M. ANHOLCER, B. BOSEK, AND J. GRYTCZUK, *Majority choosability of digraphs*, arXiv: 1608.06912, (2016).

[7] P. BALISTER, B. BOLLOBÁS, J. LEHEL, AND M. MORAYNE, *Repeated degrees in random uniform hypergraphs*, SIAM Journal on Discrete Mathematics, 27 (2013), pp. 145–154.

[8] J. BALOGH, B. BOLLOBÁS, AND R. MORRIS, *Bootstrap percolation in three dimensions*, The Annals of Probability, 37 (2009), pp. 1329–1380.

[9] S. BESSY, A. BONATO, J. JANSSEN, D. RAUTENBACH, AND E. ROSHANBIN, *Bounds on the burning number*, Discrete Applied Mathematics, 235 (2018), pp. 16 – 22.

[10] B. BOLLOBÁS, *On generalized graphs*, Acta Math. Acad. Sci. Hungar., 16 (1965), pp. 447–452.

[11] ——, *Determination of extremal graphs by using weights*, Wiss. Z. Techn. Hochsch. Ilmenau, 13 (1967), pp. 419–421.

[12] ——, *On a conjecture of Erdős, Hajnal and Moon*, Amer. Math. Monthly, 74 (1967), pp. 178–179.

[13] B. BOLLOBÁS, *Modern Graph Theory*, vol. 184 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1998.

[14] B. BOLLOBÁS AND A. SCOTT, *Independent sets and repeated degrees*, Discrete Mathematics, 170 (1997), pp. 41–49.

[15] B. BOLLOBÁS AND A. THOMASON, *Highly linked graphs*, Combinatorica, 16 (1996), pp. 313–320.

[16] A. BONATO, J. JANSSEN, AND E. ROSHANBIN, *Burning a graph as a model of social contagion*, in Algorithms and Models for the Web Graph, A. Bonato, F. C. Graham, and P. Prałat, eds., Cham, 2014, Springer International Publishing, pp. 13–22.

[17] A. BONATO AND T. LIDBETTER, *Bounds on the burning numbers of spiders and path-forests*, ArXiv e-prints, (2017).

[18] J. BOURGAIN AND L. TZAFRIRI, *Restricted invertibility of matrices and applications*, in Analysis at Urbana, Vol. II (Urbana, IL, 1986–1987), vol. 138 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1989, pp. 61–107.

[19] Y. CARO, Y. LAURI, AND C. ZARB, *Equating two maximum degrees*, ArXiv e-prints, (2017).

[20] Y. CARO, A. SHAPIRA, AND R. YUSTER, *Forcing k-repetitions in degree sequences*, Electronic Journal of Combinatorics, 16:#R7 (2013).

[21] Y. CARO AND D. WEST, *Repetition number of graphs*, Electronic Journal of Combinatorics, 16:#R7 (2009).

[22] Y. CARO AND R. YUSTER, *Large induced subgraphs with equated maximum degree*, Discrete Mathematics, 310 (2010), pp. 742–747.

[23] G. CHEN, P. ERDŐS, C. ROUSSEAU, AND R. SCHELP, *Ramsey problems involving degrees in edge coloured complete graphs of vertices belonging to monochromatic subgraphs*, European Journal of Combinatorics, 14 (1993), pp. 183 – 189.

[24] G. CHEN AND R. SCHELP, *Ramsey problems with bounded degree spread*, Combinatorics, Probabilty and Computing, 2 (1993), pp. 263 – 269.

[25] S.-Y. CHOI AND P. GUAN, *Minimum critical squarefree subgraph of a hypercube*, in Proceedings of the Thirty-Ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing, vol. 189, 2008, pp. 57–64.

[26] L. CSABA, R. FAUDREE, A. GYÁRFÁS, J. LEHEL, AND R. SCHELP, *Networks communicating for each pairing of terminals*, Networks, 22 (1992), pp. 615–626.

[27] J. CUTLER AND A. RADCLIFFE, *The maximum number of complete subgraphs in a graph with given maximum degree*, Journal of Combinatorial Theory, Series B, 104 (2014), pp. 60 – 71.

[28] P. ERDŐS, *On the number of complete subgraphs contained in certain graphs*, (1962).

[29] ——, *On theorem of rademacher-turán*, Illinois Journal of Mathematics, 6 (1962), pp. 122–127.

[30] ——, *On the number of triangles contained in certain graphs*, (1963).

[31] ——, *On the number of complete subgraphs and circuits contained in a graph*, (1968).

[32] P. Erdős, A. Hajnal, and J. W. Moon, *A problem in graph theory*, Amer. Math. Monthly, 71 (1964), pp. 1107–1110.

[33] J. R. Faudree, R. J. Faudree, and J. R. Schmitt, *A survey of minimum saturated graphs.*, Electron. J. Comb., DS19 (2011), p. 36.

[34] R. Faudree, A. Gyárfás, and J. Lehel, *Minimal path pairable graphs*, Congressus Numerantium, 88 (1992), pp. 111–128.

[35] R. J. Faudree, *Properties in path-pairable graphs*, New Zealand Journal of Mathematics, 21 (1992), pp. 91–106.

[36] R. J. Faudree, A. Gyárfás, and J. Lehel, *Path-pairable graphs*, Journal of Combinatorial Mathematics and Combinatorial Computing, 20 (1999), pp. 145–157.

[37] M. Ferrara, M. S. Jacobson, F. Pfender, and P. S. Wenger, *Graph saturation in multipartite graphs*, J. Comb., 7 (2016), pp. 1–19.

[38] S. Finbow and G. MacGillivray, *The firefighter problem: a survey of results, directions and questions*, Australasian J. Combinatorics, 43 (2009), pp. 57–78.

[39] S. L. Fitzpatrick and L. Wilm, *Burning Circulant Graphs*, ArXiv e-prints, (2017).

[40] A. Frank, *Packing paths, circuits, and cuts - a survey*, in Paths, Flows and VLSI-Layouts, B. Korte, L. Lovász, and A. S. H-J. Prömel, eds., Springer Verlag, 1990, pp. 47–100.

[41] A. M. Frieze, *Disjoint Paths in Expander Graphs via Random Walks: a Short Survey*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1998, pp. 1–14.

[42] Z. Füredi, *Graphs with maximum number of star-forests*, Studia Scientarium Mathematicarum Hungarica, 27 (1992), pp. 403 – 407.

[43] A. GIRÃO, T. KITTIPASSORN, AND K. POPIELARZ, *Generalized majority colourings of digraphs*, Combinatorics, Probability and Computing, 26 (2017), p. 850âĂŞ855.

[44] A. GIRÃO, G. MESZAROS, K. POPIELARZ, AND R. SNYDER, *On the maximum diameter of path-pairable graphs*, ArXiv e-prints, (2017).

[45] A. GIRÃO AND K. POPIELARZ, *Large induced subgraphs with k vertices of almost maximum degree*, ArXiv e-prints, (2017).

[46] A. GIRÃO, T. KITTIPASSORN, AND K. POPIELARZ, *Partite saturation of complete graphs*, ArXiv e-prints, (2017).

[47] A. GIRÃO, G. MÉSZÁROS, K. POPIELARZ, AND R. SNYDER, *On the maximum degree of path-pairable planar graphs*, ArXiv e-prints, (2017).

[48] E. GYŐRI, T. R. MEZEI, AND G. MÉSZÁROS, *Note on terminal-pairability in complete grid graphs*, Discrete Mathematics, p. appearing.

[49] ——, *Terminal pairability in complete graphs*, Journal of Combinatorial Mathematics and Combinatorial Computing, pp. submitted, arXiv:1605.05857.

[50] A. HAJNAL, *A theorem on k-saturated graphs*, Canad. J. Math., 17 (1965), pp. 720–724.

[51] T. HIRATA, K. KUBOTA, AND O. SAITO, *A sufficient condition for a graph to be weakly k-linked*, Journal of Combinatorial Theory, Series B, 36 (1984), pp. 85 – 94.

[52] A. HUCK, *A sufficient condition for graphs to be weaklyk-linked*, Graphs and Combinatorics, 7 (1991), pp. 323–351.

[53] J. R. JOHNSON AND T. PINTO, *Saturated subgraphs of the hypercube*, Combin. Probab. Comput., 26 (2017), pp. 52–67.

[54] L. KÁSZONYI AND Z. TUZA, *Saturated graphs with minimal number of edges*, J. Graph Theory, 10 (1986), pp. 203–210.

[55] K. KAWARABAYASHI, Y. KOBAYASHI, AND B. REED, *The disjoint paths problem in quadratic time*, Journal of Combinatorial Theory, Series B, 102 (2012), pp. 424 – 435.

[56] T. KITTIPASSORN AND G. MÃĽSZÃĄROS, *Frustrated triangles*, Discrete Mathematics, 338 (2015), pp. 2363 – 2373.

[57] F. KNOX AND R. ŠÁMAL, *Linear Bound for Majority Colourings of Digraphs*, arXiv: 1701.05715, (2017).

[58] D. KORÁNDI AND B. SUDAKOV, *Saturation in random graphs*, ArXiv e-prints, (2015).

[59] S. KREUTZER, S.-I. OUM, P. SEYMOUR, D. VAN DER ZYPEN, AND D. R. WOOD, *Majority Colourings of Digraphs*, arXiv: 1608.03040, (2016).

[60] E. KUBICKA, G. KUBICKI, AND J. LEHEL, *Path-pairable property for complete grids*, Combinatorics, Graph Theory, and Algorithms, 1 (1999), pp. 577–586.

[61] M. R. LAND AND L. LU, *An upper bound on the burning number of graphs*, in Algorithms and Models for the Web Graph, A. Bonato, F. C. Graham, and P. Prałat, eds., Cham, 2016, Springer International Publishing, pp. 1–8.

[62] D. G. LARMAN, *On the number of complete subgraphs and circuits in a graph*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 308 (1969), pp. pp. 327–342.

[63] R. LIPTON AND R. TARJAN, *A separator theorem for planar graph*, SIAM Journal on Applied Mathematics, 36 (1979), pp. 177–189.

[64] L. LOVÁSZ AND M. SIMONOVITS, *On the number of complete subgraphs of a graph ii*, (1983), pp. 459–495.

[65] G. MÉSZÁROS, *Note on the diameter of path-pairable graphs*, Discrete Mathematics 337 (2014), p. 83âĂŞ86.

[66] ——, *On path-pairability in the cartesian product of graphs*, Discussiones Mathematicae Graph Theory 36 (2016), pp. 743–758.

[67] M. MIDDENDORF AND F. PFEIFFER, *On the complexity of the disjoint paths problem*, Combinatorica, 13 (1993), pp. 97–107.

[68] D. MITSCHE, P. PRAŁAT, AND E. ROSHANBIN, *Burning graphs: A probabilistic perspective*, Graphs and Combinatorics, 33 (2017), pp. 449–471.

[69] N. MORRISON, J. A. NOEL, AND A. SCOTT, *Saturation in the hypercube and bootstrap percolation*, Combin. Probab. Comput., 26 (2017), pp. 78–98.

[70] V. NIKIFOROV, *The number of cliques in graphs of given order and size*, Transactions of the American Mathematical Society, 363 (2011), pp. 1599–1618.

[71] E. NORDHAUS AND B. STEWART, *Triangles in ordinary graphs*, Canadian Journal of Mathematics, (1963), pp. 33–41.

[72] B. ROBERTS, *Partite saturation problems*, Journal of Graph Theory, 85 (2017), pp. 429–445.

[73] N. ROBERTSON AND P. SEYMOUR, *Graph minors. XIII. The disjoint paths problem*, Journal of Combinatorial Theory, Series B, 63 (1995), pp. 65 – 110.

[74] K. A. SIM, T. S. TAN, AND K. B. WONG, *On the burning number of generalized petersen graphs*, Bulletin of the Malaysian Mathematical Sciences Society, (2017).

[75] E. SULLIVAN AND P. S. WENGER, *Saturation numbers in tripartite graphs*, Journal of Graph Theory, 84 (2017), pp. 428–442.

[76] R. THOMAS AND P. WOLLAN, *An improved linear edge bound for graph linkages*, European Journal of Combinatorics, 26 (2005), pp. 309 – 324. Topological Graph Theory and Graph Minors, second issue.

[77] J. VYGEN, *Disjoint paths*, Report No. 94816-OR, Research Institute for Discrete Mathematics, University of Bonn, (1994).

[78] W. WESSEL, *über eine Klasse paarer Graphen. I. Beweis einer Vermutung von Erdős, Hajnal und Moon*, Wiss. Z. Techn. Hochsch. Ilmenau, 12 (1966), pp. 253–256.

[79] ——, *über eine Klasse paarer Graphen. II. Bestimmung der Minimalgraphen*, Wiss. Z. Techn. Hochsch. Ilmenau, 13 (1967), pp. 423–426.

[80] A. A. ZYKOV, *On some properties of linear complexes*, Mat. Sbornik N.S., 24(66) (1949), pp. 163–188.